

```
Db 21 NLDVBSPIVFPDAA8FGQTVVQGGSRLLVVGAPLEAVNQTGRLYDCAPATGMCQPIV 80
Qy 62 IQVPVEAVNMSIGLSLAATTSPPQLLACGPTVHTQCTSENTYVVKGLCFLFGSNLRQOPKF 121
Db 81 LRSPLRAVNMSIGLSLWATNNAQLACGPTAQRACVQWYAKGSCULLGSSSL-QFIQAV 139
Qy 122 BEALRGCPQSDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEFR 181
Db 140 PASMECPPEMDIAFLIDSGSINORDFAOMDFKALMGFASTSTFLSMOYENILK 199
Qy 182 IHTFKEFQNNPNRSLIKPITOLLGRTHATGLRKVVRLFNITNGARKNAFKILFLT 241
Db 200 THFTTFEKNLDPQSLVDPIVOLQGLTYTATGIRTWMEELFESKGRSAKILLVIT 259
Qy 242 DGBKFGDPLGYEDVIPDLREGVIRVIGVGDAPRSEKSRQELNIVASKPRPRHVTQINN 301
Db 260 DGQKYEDPLEYSDVIPADKAGHIRVAIGVGDAPQEPALKELTIGSAPPQDHVFKVGN 319
Qy 302 FEALKTIONLREKI FAIEGTQCTSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAGG 361
Db 320 FAALRSIQRLQOEKI FAIEGTQSRSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAGG 379
Qy 362 VFLYTSKEKSTFNTMTVDSDMDAYLGAAAILNRVOSLVLGAPRYQHI GLVAMFRQ 421
Db 380 AFLYPNTRPTFINMSQENVDRSDYLGYSTAVAFWKGVHSLILGAPRHQHTCKWIFTQ 439
Qy 422 NTGMWESNANVKTQICAYFGASICSVDVDSNGSTDILVLIGAPHYVQTRGGVSVCLP 481
Db 440 EAHWRKSEVRGTQIGSYFGASLCSVDVDRGSTDILVLIGAPHYVQTRGGVSVFPVP 499
Qy 482 RGQRAQCDAVLVYGEQGWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLFH 541
Db 500 -GVGRQWQCEATLHGEQCHPWGRGVALTVLGVNGDNLADVAIGAPGESRGAIVYFH 558
Qy 542 GTSOGSISPHSQRISAKSLSPLOYGQSLSGQDLTMGLVDTLTVGACHVILLERSQP 601
Db 559 GASLLEIMPSPQSVRTQSLRLQYFGQSLSGQDLTDGLVDVLAVGACHVILLERSLP 618
Qy 602 VLRVKALMESNPREVAARVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQTQSVVTV 661
Db 619 LKVELSIRAPAPMEVAKAVQCWERTPTVLEAGATVCLTVHKGSPDLL-GNVQGSVRY 676
Qy 662 DLALDSRGRPSRAVFNETHKSTRQTQVLGHTQTCETLKLQPLNCIEDPSPVILRLNPS 721
Db 677 DLALDPGLISRAIPEDETQKNTLGRKTLGEGDCHETVKLLPDCVEDAVSPILRLNFS 736
Qy 722 LVGTPLAGNLRVLAEDAQLFTALPPFEKNGCNDNI CODLSITFFSFMSLDCLVVG 781
Db 737 LVRSASAP-RNLHPLVAVGSQDHITASLPFEKCKQELLCEGDLGIFNFSGLQVLVVG 795
Qy 782 PRFNVTVVRNDEGDSYRTQVTPFPFLDLSYRKVSTLQNRQSRWRLLACESASSTVS 841
Db 796 SPFLTIVTVVWNEGEDSYGLVAVFYAGLSYRVTGTQ-QPHQYPLRLACEAPAAQED 854
Qy 842 GALKSTCSINHPIFPENSEVTENITPDVDSKASLGKLLKANVTSENNNPKNTKTEFO 901
Db 855 --LSSSCSINHPIFREKAKTITMITFDVSYKAFGLGRLLLRKASSENKXPTNKTAFO 912
Qy 902 LELPKVAVYVMTVSHGYSKYLNTASENTSR-VMQHYQVSNLXGORSIPISLVLPVP 960
Db 913 LELPKVYTVYTLISQRODSTNHNFSSSHGRRQEAARVYVNNLSPLKLAVRVNFVWPV 972
Qy 961 PLNQTVIWRDQVTFESENLS--TCHTKERLPSHSDFLAELRKAPVNVCSLAVOCQIOCDI 1019
Db 973 LLNGVAVMD----VTLSSPAQGVCSVQKHPQNPQDFITQIRRVLDCCSADCLHFRCDI 1029
Qy 1020 PFFGQIBEFNATLKNLSFDWYIKTSHNHLIVASTAILPNDVSFTLLPGQAFVRSQTE 1079
Db 1030 PSLDIQDELFDILRNLNSFGVSVQTLQEKVLLVSEABITFTDSVYSLPGQEAFLRAQVE 1089
Qy 1080 TKVZPFEVNPDLPLVGVSSVGGILLALLITAAVLKIGLGFKKQYKMM 1126
Db 1090 TTLSEYVYVYFIFLVAGSSVGGILLALLITVVLVYKLGFFFRQYKEML 1136
```

RESULT 7

Q9WTV4

```
ID Q9WTV4 PRELIMINARY; PRT: 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSP; P20701; ILPA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;
```

Query Match 25.9%; Score 1523.5; DB 11; Length 1161;

Best Local Similarity 33.9%; Pred. No. 4.5e-103;

Matches 396; Conservative 220; Mismatches 461; Indels 91; Gaps 37;

Qy 1 FNLDTENAMTFQENA-RGFGQSVVOL-QGSRVVVVGQELVAANQORSLVQCDYSTGSCBP 59

Db 24 YNLDTPTQSLAQAGRHFGYQVLQEDG-VVVGAPGE--GNTGGLHCRSTSSFCQP 79

Qy 60 IRLQVPEAVNMSIGLSLAATTSPPQLLACGPTVHTQCTSENTYVVKGLCFLFGSNLRQOPQ 119

Db 80 VSLH-GSNHSTKYLGWTLATDAAGKSLACDPGLSRTCDQNTVLSGLCYLFPQSLGPM 138

Qy 120 KPPEALRGCPQSDSDIAFLIDSGSIIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEE 179

Db 139 QNRPAYQECMKGVLDVLPFLDGSQSLDRKDFEKLLEFMKDVNRKJSTNTSYQFAAVQFSTD 198

Qy 180 FRTHFTFKCF-QNNPNRSLIKPITOLLGRTHATGLRKVVRLFNITNGARKNAFKILF 238

Db 199 CRTEFTFLVYKQKNPDLVLSGVQPMFLTTNTFRAINVVAHVFKESGARDPAKVLV 258

Qy 239 LFTDGEKF--GDLPGYEDVIPDLREGVIRVIGVGDAPRSEKSRQELNIVASKPRDHV 296

Db 259 IITDGEASDKENISAHD-----ITRYIIGIKGFVSQKTLHIFASEPVEFV 309

Qy 297 FQINNFEALKTIONLREKI PAIEGTQCTSSSFHEMSQEGFSAATNSGPELLSTVGSY 356

Db 310 KILDTTEKLDLPTDLQRIYAIIEGTRNQLDYSFNNELSSGSIADLSKGHAVVGAVGAK 369

Qy 357 DWAGGVF-LYTSKEKSTFFINMTRVSDMDNDAYLGAAA-IILNRVQSLVILGAPRYQHIG 414

Db 370 DWAGGFLDLREDLQAGTFVQGEPLTSDVRGVGLGYTVVAVMTSRSSRPLLAAGAPRYQHV 429

Qy 415 LVAMFR-QNTGMWESNANVKTQICAYFGASICSVDVDSNGSTDILVLIGAPHYVQTRG 472

Db 430 QVLLFQAPAEAGGRNQTKTEGTQIGSYFGSEELCSVDLDDQGBAEILLIGAPLFFGBQRG 489
 Qy 473 GOVSVCPLPQGRQARWQCDLVLGEQOPQGRFGAALTVLGDVNGDKLTQVAIGAPEED 532
 Db 490 GRVPTY---QRRQSLFEMWSELQDPCYPLGRFGAAITALTIDINGDRLTOVAVCAPLEE- 545
 Qy 533 NRGAVYLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGGLVOLTVCAGQ 592
 Db 546 -QGAVIYFNQKPG-GLSPQPSQRQGAQVFGIRWFGRSIHGVKDLGGDLADVVVGPBG 603
 Qy 593 HVLLRSQPVLVRKAIINFEPREVARNVFENDQVVKGEAG-EVRVCLHVQKSTRDLRL 651
 Db 604 RVVLSRPVVDVTVLSPSEPIPVHEVECSYAREEQHGKVKACFAIKPLTPQ--F 661
 Qy 652 EGQIQSVVTVDLALDSORPHSAVFNETKASTRAQTQVLGTQTCETLKLQLPNCIBDPV 711
 Db 662 QGRLLANLSTQLQDGRHMSRGLFPDGSHELSGNTSITP-DKSLDLPFHFPFICIQDLI 720
 Qy 712 SPIVLRNLFSLV---GTPLSAFGN-LRPVLAEADQRLFTALFFPEKNCMDNTICQDLSI 767
 Db 721 SPINVSINFLSLLBEGCTPROKGRAMQPILRPSHTV-TKEIFPEKNCGEDKCEANLTL 779
 Qy 768 TFSFMSLDCLVVGQP-----REFNVTVTVRNDCEDSYRTQVTFPFLDLSYRKVSTLQN 821
 Db 780 SSPARS-----GPLRMSSASLAVSWTLNSGEDAYTVRLDLDLDFPRGLSFRKVEMLQ- 831
 Qy 822 QRSQSWRLACESASSTEVSGAL-KSYSCSINHPIPFENSESEVTENITFDVDSKASLGNKL 880
 Db 832 --PHSRMPVSCBEL--TEGSSLLFKTKLKNVSSPIFQAGQSVSLQVNFNILLNSWEDFV 887
 Qy 881 LLKANVTSEN-NMPTNKTEFLQELPVRKYAVYVWTVSHGVSTKYLNFTASENTSRVMQHQ 939
 Db 888 ELNGTVHCENENSSLOEDNSAATHIPVLYPWNILTKQENSTLYISPTPKGPKTQQVQHV 947
 Qy 940 YQV---SNLQORSIPISLIVFLVRLNQTVYNDRPQ--VTFSENLS-----TCHTKB-RL 989
 Db 948 YQVRIQPSAYDHNMP-TLEALVGP-----WPHSEDPITYTVWSVQTDPLVTVCHSEDLKR 1000
 Qy 990 PSHSDPFLAELRKAPVWNCIAVCCORIQCIDIPFGIQBEFNATLKGNSLFDWYIKTSHNL 1049
 Db 1001 PSSE--AEQCLPGV-----QPRCPVP--RREILQVTVTVLSKEIKAS-STL 1045
 Qy 1050 LIVSTAILFNDSVFTLLPGGAVRQGTETKYVFEFVNPNDPLIVGSSVGGLLIALIT 1109
 Db 1046 SLCSLSVSFNSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLRVYVLSGTGGLVLLFLF 1104
 Qy 1110 AALVKLGFFKQYKDMX-SEGPGPGAEP 1136
 Db 1105 LALYKVGFFKRNLEKEMADGGVYNGSP 1132
 RESULT 8
 Q9R200 PRELIMINARY; PRT; 1160 AA.
 AC Q9R200
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Integrin alpha L.
 GN ITGAL.
 OS Mus musculus (Mouse),
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spleen;
 RA Ma R.Z., Teuscher C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF05901; AAD25884.1; --
 DR HSSP; P20701; ILFA.
 DR MGD; MGI:96606; Itgal.
 DR GO; GO:0008395; C:integrin complex; IEA.


```

Db 780 SSPARS-----CPLRLSSASLAVZTWLNSGEDAYVRLDLDPRGLSRFKVEMLO- 831
Qy 822 QRSORSWRLACASASSTEVSGAL-KSTSCSINHPHPENSEVTFNITFDVDSKASLONKL 880
Db 832 --PHSRMPVSCHEL--TEGSLTKTKLCNVSSPIFKAGQEVSLQVNFNTLMSWEDFV 887
Qy 881 LLKXANTSEN-NMPTNKTETFOELPKVAYVMVTSHTKYLNFPTASENTRVMQHQ 939
Db 888 ELNGTVCHEMENSLOEDNSAATHIPVLPVNLTKXENSTLYISFTPKGPKTQOVHV 947
Qy 940 YQVSNLQORSPLSLVFLVRLNQTVINDRPQ-----VTFSEMSS-----TCHKE-RLP 990
Db 948 YQV-----RIQPSAYDNMPT-LEALGVVERPHSEDLITYTWSVQTDPLTCHSEDLKRP 1001
Qy 991 SHSDFLAELRKAPVNVCSIAVCORIOCDIPFGIOBEFNATLKNLSFDVYIKTSENHLL 1050
Db 1002 S-----SEAPCLPGV--QFRCPVIF--RWEILQVGTVELSKHETAS-STLS 1045
Qy 1051 IVSTAELFNDVSFTLLPGQAFVRSQTEKVPFFVFNPLPIVGVSSVGLLLALITA 1110
Db 1046 LCSSLVSFNSSKHFHLYGSKA-SEAQVLVKVDLHEKEMLVVYVLSGIGLVLLFLFL 1104
Qy 1111 ALYKLGPFKEQYQDMW-SEGPPCAEP 1136
Db 1105 ALYKVGPFKRNLEKMEADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY; PRT; 1196 AA.
AC Q98TF1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD11-1.
GN CYA1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=962;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048536; BAB39134.1; -.
DR HSSP; P20701; 1LPA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Approtease A5.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807B7DCA53B CRC64;

Query Match 23.7%; Score 1395; DB 13; Length 1196;
Best Local Similarity 31.2%; Pred. No. 1.5e-93;
Matches 371; Conservative 227; Mismatches 448; Indels 144; Gaps 39;
Qy 1 PNLTEHMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQEIYAANQSGSLYQCDISTGS 56

```

```

Db 32 FNIDTEHPLRFNCAPEDPFGYSVYQTEFGNRKQIIVGAPLE---GNSTGEWYCTADLQS 88
Qy 57 CDPRLQVP-----VRVNMISGLSLAATSPQQLACGPTVHTQTSENTYVYVGLCFPLFS 112
Db 89 CQ--BLQRPGSSEVRFPFGMSAAVSSAALTS-----CSFYPFHECDGNSYLVGVCYQFS 140
Qy 113 NLRCQPKQFPBALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKEWVSTVMEQKLSKTLFS 172
Db 141 SL-QAVSNPTAAYQCSKREVNVLVFLDGSSSMKTVDFEMNENFKIDIMKLSNSIKFA 199
Qy 173 LMQYSEEPRIHTPKFQNNPNRSLIKPI TOLLGRTHATGLRKVVRLEFN-ITNGARK 231
Db 200 AVQFSTVTRTDFPDNDYQSGSAEELMKE-TMKSLTNTTHKAIIDYILKLLNSMLSGADS 258
Qy 232 NAFKILPLLTCGEKEDPLGYED--VIPELDREGVIRVVGVDAPFRSKSQEINTVAS 289
Db 259 KAKALVIITD---GDFSDDDYVNLKCDQNIIRVIIVG-----KVDLIELTQLAS 309
Qy 290 KPRDRHVFOINNFEALKTIONQREKIPAEICTQTGSSSPHEHMSQEGFSAITSNGPL 349
Db 310 EPKRNTTFYIKDYSGLKGLLDNLQKKIYNIEGSDAQGRDRQKLSQSGFSVYOBESVI 369
Qy 350 LSTVGSYDAGGVFLYTSKEKSTFINMTRVDSDMN-DAVLGYAALILNRRVQSVLIGAP 408
Db 370 VGSVGSNDWRGALYEVTG-SGSDFKETIIDPAVNKDSYMGYSTVLGMRGVSLLFSGAP 428
Qy 409 RYQHIGLVAMFRQNTCMMESNANVRGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYTE 468
Db 429 RAHTGLVTLFTKNQNTWTVTNSINGEQIGSYFGASLSLLDSDSDSDFLLVGLAPLPYQ 488
Qy 469 QTRGQSVCPPLPGQRAKWCDAVLYGEQ-----GQPMGRFGAALTVLGDVNGDKLTIV 523
Db 489 SQ-----PRTEGLYVYLSSEQYFKTLNVQSSTGRFAASVASLKDNGDGLSDV 540
Qy 524 AIGAGEHDNRGAVYLFHGTSGSGSPSHS-ORIAAGSKLSPLRLQVFGQSLSGQDLTMDG 582
Db 541 AVGAPLE--NEGVIYILGDRTHGINPLTQRIISVQSVLPGLOQFGVSLTGQMDWNNN 598
Qy 583 LVDLTVAGAGHVLLRSQPLVKALMEENPREVARNPECDQVVKGEAEVRLVAV 642
Db 599 LTDIVIGAQGIIVLKAREFVMSVAQLSFPSEISLNYFECFSG--NAFNALNLTSCFTV 656
Qy 643 QKSTRDLREGQIQSV--VYDIALDSDGRPHRAVFNETKNSTR--RQTVQLGLTQTCET 698
Db 657 TERTSS---TGSLEKLVNLVNLVNVVGRMGRGFPDQSSVSRTLQQSVLLDSGSSCFN 713
Qy 699 LKQLPNCIEDPVSPFIVLRLNFSLVGTPLSAGNLRPLVLAEDAQRLLFTALPPEKMGND 758
Db 714 FSIFMLRCVADTVSPLKIRMFNS--QTEMLS-GNSVAVLDVHSRTBEENVEVFORCNSN 770
Qy 759 NLCQDDLSITPFSMSLDCLVGGRPRENVTVTVRDGDSYRTQVTFPPPLDLVYRKVST 818
Db 771 NSCVADLKUNFSTN-NTLVVENQAHFTVQVSLANPGDDSYNTSVLHVPEGISLXKPA 829
Qy 819 LQNRQSQRSWRLACHSASSTEVSGALKSTCSINHPFIPENSEVTFNITFDV---DSKAS 875
Db 830 IKPSRTR-----SSCGDRDSGATNRTTCSIDLPLVYRSGTTPOTPLGTFRWKMDNDS 881
Qy 876 LGVLLKLVKANVTSENNKPTNKTETFOELPKVAYVMVTSHTGV-STKYLNFPTASENTR 934
Db 882 NRMETIMTANSNNGNM---SDTEVRSVVFQVADLALSILVAEDSVTTFNFSLEDRGPK 938
Qy 935 VMQHOYQVSNLQORSPLSLVFLVPR-----LNTFPMSEHVHHSFISSYHQIIMCLN 963
Db 939 ELNITYKVENSGKDLFVSVTLTLPCQPHVILTPTFPMSEHVHHSFISSYHQIIMCLN 998
Qy 964 QTVVDRPQVTESEMLSSCTCHKEKRLPSSHDFLAELRKAPVNVCSIAVQRIQ-----C 1017
Db 999 KHLFPFSLSAVQQVQVSTGWSLREVEC-SQP--DLNKSVAHFNLADARLQNVKEYS 1055
Qy 1018 DIPFFGICQEB--FNATLKNLSFDVYIKTSENHLLIVSTAELFNDVSFTLLPGQAFVR 1075

```

```

Db 1056 KVSFYEFRKDXNVFSISAEIYNLSVNTS-----SELKYNPH-----R 1094
Qy 1076 SQTETKVEFEVFNPEPLIV-GSSVCGLLLLALITAAALYKLGFFKRYKXD 1124
Db 1095 SOTEKVKV-FVVPSPSLMLIVCTGAUGGVFFLLIILFLKLCGFFKRNRED 1143

RESULT 10
O96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC068777; AA08777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6FF2546B8C632F9 CRC64;

Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.1%; Pred. No. 2.4e-30;
Matches 375; Conservative 187; Mismatches 432; Indels 175; Gaps 37;

Qy 1 FNLDTENAMTFQ--ENARFGQSVWLOGSRVVVGAPQEIIVAAORGSIYQCDYSTGSC 58
Db 26 YNLDRVGRSFPSPRAGRHFGRVLOV-GKGVIVGAPGE---GNSTGSLYQCQSGTGCL 81
Qy 59 PLELOVPEAVNMSLGLSLAATSPQLLACGFTVCTSENTYVVKGLCFLGSLNRQOP 118
Db 82 EVTLR-GSNYTSKYLGMTLA--TDP----- 103
Qy 119 QKPEALRGCPQSDSIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSE 178
Db 104 -----TDGS-----ILFAAQVQST 117
Qy 179 EFRHITFEFQNPNPRLSIKPIITOLLGTHTATGLRKVVRLNITNGARKNAKILP 238
Db 118 SYKTEPDSFYVRKDPDALLKGVKMLLTNTFGAINVYATVFPREELGARPDATKVL 177
Qy 239 LLTDGKFGDPLGVEDVIPLEDRGVIRYVIGVDAPRSEKSRQELNLTVAKPPDRHVFQ 298
Db 178 IITDGE--ATDSGNIDAADK-----IIRVIIGIKHFQKESQETLHKFASKPASEFKI 230
Qy 299 INNFPAKTIQNLRKIPFAIEGTQTCGSSSPHEMSQBGFSAAITNSGPIILSTVGSYDW 358
Db 231 LDTFEKLKDLFTBLQKIYVIEGTSLQDLTSFNNELSSSGISADLSRGHAUVGAVGAKDW 290
Qy 359 AGGVF-LYTSKSKSTFTNTRVSDMNDVAVGLYAAA-IILRNVRQSLVLGAPRYQHIGLV 416
Db 291 AGGFLDLKADLQDDTTFIGNEPLEPTEVRAGVLGTVTLWLSRQKTSLLASGAPRYQEMGRV 350

```

```

Qy 417 AMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQ 474
Db 351 LLFREPQGGHWSQVTHGTQIGSYFGGELCGVDVDQDGETELLIGAPFYGEQGR 410
Qy 475 VSVCLPRGQARMQCDVAV--LYGEQGPWRGFGAALTVLGDVNGDKLTDVAIGAPGED 532
Db 411 VFYI-----QRRQLGFEVSELOQDGPYPLGRFGEAITALTIDINGDGLVDVAVGAPLEB- 464
Qy 533 NRGAVYLPHEGTSGSGISPSHSQRIAGSKLSFELQVFGOSLGGQDLTWDGLDVLTVGAQG 592
Db 465 -QGAVYIFNGRHG-GLSPQSQRIEGTQVLSGIQWFGRSIHGVKDLLEGGLADVAVGAE 522
Qy 593 HVLLRSQVLRVKAIMEFNPFEARNVFECDQV--RGKEAGEVRVRLHVQKSTRDLRL 651
Db 523 QMIVLSRPPVDMVTLMSFSPAIPVHEVECSYSTSNKMKGVNITTCFQI-KSLIPQF- 580
Qy 652 EGQIQSVVYDLDLSDSGRPHSAVNETKNSRGTQVLTQCTETLKLQPCNIEBPV 711
Db 581 QGRLVANLTYTLQDLGHTRRRGLPFGGRHLLRNIAVT--TSMCTDTSFFHFPVVCQDLI 639
Qy 712 SPIVLRIMFSL---VGTPLSAFEN-----LRPVLAEQAQLFTALFPFKKNCGNNDICQ 762
Db 640 SPINVSINFLSWEBSGTTRDQORAGKDIPILRPSLSHSETWEI-----PFEXKCGEDKKCE 694
Qy 763 DDLSTTFMSLDCLVVGCPREFVNTVTRNDGEDSYRTQVTFPPPLDLRYKXVTLQNO 822
Db 695 ANLRVSFSPARSRALRLTAPASLSVELSLNLEEDAYVQLDLHPFPGLSFKVEML--- 751
Qy 823 RSQSRWRLACES--ASSTEVSGALKSTCSINHPIFFENSEVTFNITPDVDSKASLGNKL 880
Db 752 KPHSQIPVSCHELPEESRLSRAL---SCNVSSPIFKAGHSVALQMMNTLVNSWGDV 808
Qy 881 LKANVTSTENN---MPTNKTEFOLELPVYAVVMVTSHGVSVKYLNFTASENTRVM 936
Db 809 ELHANVTGNEDSDLEDNSATTI---IPILYPINILIQDOEDSTLYVSFTPKGKIHOV 865
Qy 937 QHQYOV---SNLGQSRSLP-ISLVFLAVPVLQVLTWDRPQVTFSENLASTCHTK--ERLP 990
Db 866 KHYQVRIQPSIHDNIPTLRAVGVQPPSPSGPITHQWSVQMEPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVNVNCSIAVCQRIQCDIPFFGIGIEFNATLKNLSFDWYIKTSHNH 1048
Db 924 DAAEPCLPCALFRCPVW-----FRQELVQVIGTGLVGEIAS-SM 964
Qy 1049 LLIVSTAILFNDVSFTLLPGCGAFVRSQYETKVPPEVPNPPLIVGSSVGLLLALI 1108
Db 965 FSLCSLSLSTSFNSKHFHLYGSNASL-AQVWVKVDVYVEKQKMLYLVLSGIGGLLLLI 1023
Qy 1109 TAALYKLGFPKRYKQKMMSEG-GPPGAP 1136
Db 1024 FIVLYKVGFFKENLKEKMEAGRGVNGIP 1052

RESULT 11
Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD11-2.
OS Cyprinus carpio (Common carp).
GN CIA2.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";

```


Db 489 -BEDAFVRRLEGQMGSGYFGSVLCPVDIDNDGTTDFLLVAAPFYHIRGBEGRVYVQVPE 547
Qy 483 GQARWQCDAVLVGEQGWGRGAALTIVLGVNGDKLTDVAIGAP-----GEEDNRGA 536
Db 548 -QDASFSLAHTLSGHPLNRSRFGFMAAVGNDINQDKFTDVAIGAPLEGFGAGDGASYS 606
Qy 537 VYLFHGTSGSGISPSQRSIAGSKLSPRLQYFQSLSGGQDLTMDGLVLTIVGAQGHVLL 596
Db 607 VYIYNGHSG-GLYDPSQIRASSVASGLHYFGMSVSGGLDPNGDGLAITVGRSDSAV 565
Qy 597 LRSQPLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKS---TRDLRREG 653
Db 666 LRSRPVVDLTVMSTFT?-----DALEMFVFIKRY--DVNLCFEVDSSWASEPGLRM 715
Qy 654 QIQSVVTVYDLALDSCGRHSPAVNEFKNSRTRTOVLGLTQTC-----696
Db 716 FLNFTVDVTV-----TKQRQLQCESSGSCQCLRWNGSGFLCEHFMLJ 760
Qy 697 ETILKQLPNCIEDPVPSPVILRNFSIVGTPLSAGNLR-----PVLAEDAQRLFTALF-P 750
Db 761 STEEL-----CEEDCFSNITIKVYE-----FQTSQGRRDYFNETL--DHYKEPSAIFQLP 809
Qy 751 FEKNCGNDNICQDLISITFSNLSGLVGVGPREFNVTVTVRNDGEDSYRTQVTFEPLD 810
Db 810 YESKCKNKVFCIAEIQIOTN-ISQQLVGVTVKEVTMNLISLTSNGSDSYTMNMLNPRN 868
Qy 811 LSYRKVSTLQNSORSQSWRLACSSASTESVGALKSTCSINHPIPFENSEVTFNITFDV 870
Db 869 LQFKKI-----QKPSVDPQCDPXPV--ASVLVWCKIGHPIL-KSSVNVSVTMQL 918
Qy 871 DSXASLGNKLLKANTSENMMPTNKTBFQLELPVKYAVVMVTVSHGVSTKYLNETASE 930
Db 919 EESVFPNRTADITVTISNKSNEKSLARETR--SLQFRHAFIAYLSR--PSVMYMN--TSQ 971
Qy 931 NTSRVNMOHOVENLQORSUPISLVFLVPLVRLNQTWDRPQVTFPENLSST-----CHT 985
Db 972 SPDHKEFPNVGENLFGAVFQLOICVPIKLQDP-----QIVRVKNLTKQDTECTQ 1025
Qy 986 KERLPSSDFLABLRKAPVNVCSIAVCORIQCDDIPFGIOEBFNATLKNLSPDWIKTS 1045
Db 1026 SOEPACGSDPVQVHWSHWSVVCAL-----TSNKENVTVAABISVG 1065
Qy 1046 HNHLLIVSTA-----RILPNDVSFTLLPGQAFVRSQJTBKVEPE-----EVENPLPLIV 1095
Db 1066 HTQLQLRDVSELPIELISFNKSLYEGLNNE-----NERTKITVFLKEEETRSPLII 1119
Qy 1096 GSSVGGILLALITALYKLGFPKRYQKDMWSE 1128
Db 1120 GSSIGGLLVLLWIIAILFKGFRKRYQOQNLME 1152

RESULT 14

O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DF 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.,"
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSSP; P11215; 1BHQ.

GO; GO:0008305; C:integrin complex; IEA.

GO; GO:0004895; P:cell adhesion receptor activity; IEA.

GO; GO:0007160; P:cell-matrix adhesion; IEA.

InterPro; IPR000413; Integrin_alpha.

Pfam; PF01839; FG-GAP; 3.

Pfam; PF00357; Integrin_A; 1.

Pfam; PF00092; vwa; 1.

PRINTS; PR01185; INTEGRINA.

PRINTS; PR00453; VWFADOMAIN.

SMART; SM00191; Int_alpha; 4.

SMART; SM00327; VWA; 1.

PROSITE; PS00242; INTEGRIN_ALPHA; 1.

PROSITE; PS50234; VWEA; 1.

FT NON TER

SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E1B95CBCEC CRC64;

Query Match 18.8%; Score 1109; DB 11; Length 1167;

Best Local Similarity 27.5%; Pred. No. 1.9e-72;

Matches 332; Conservative 206; Mismatches 42; Indels 244; Gaps 37;

Qy 47 LYQCDYSTGSCPTIRLQVPVBAVNMVSLGLSLAATTPSPOLLACGPTVH-----94

Db 67 LYQC-----ATSLIPEIGCCQVVEHLMPKGRYQGV 98

Qy 95 -----CTCSENTVVKGLCFELFQSNLRQOKPEALRG-----127

Db 99 LVRNHNGVLVCIQVQSRKPRSLNSELTCACSLTLPNLDLQAPFSDLEGVLDLGASVNS 158

Qy 128 -----CPOED-----SDIAFLIDGSGSIIPHDIFRMKKEWV 157

Db 159 GDIYSGKSGTGEETKSARRLRQAVBEDEEEAGTEALIVLDGSGSIPEPSFOKAKOFI 218

Qy 158 STVMEQL--KSKTLPFLSMQYSEBPRIHFTFKFQNNPNRSLIKPIITOLLORTATGL 215

Db 219 STMENFYEKCFECFALVQGVVQIOTFEFLDLSRDINASLAKVQSIQVQKEVTTASAM 278

Qy 216 RKVPEELFNIINGARKNAFKILFLLTDCGEKGDPLGYEDVPELDRGVIRVVGDAF 275

Db 279 QVLDNIITPGRSGRRKALKVMVLTGDIIFRDLNLTIVISSKMQGVVAFALGVGNAP 338

Qy 276 RSEKSRQELNIVASKPRDRHVQINNFEALKTIOQLREKIPATRGTOGSSSSFEHEWS 335

Db 339 ENNNTYRELKLIASDPKAAHTFKVTYSALDGLLSKLQRIIHMEGT---VGDITLYQLA 395

Qy 336 QEGPSAAITSNP--LLSTVGSVDNAGGVFLY--TSKEKSTFNTMTFVDSMDNA---YLGY 390

Db 396 QTGFSAQILDKGQVLLGTVGAFNWSGALLYNTONGRGRFLNQT--AKEDFRAAQYSLGY 454

Qy 391 AAAIILRNVRQSLVGLAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYGASLCSVDV 450

Db 455 SVAALHKAHGSVYVAGAPRHLRGAVPELQKEDGETPMRRIEGBQMSYFGSVLCPVDI 514

Qy 451 DNGSTDLVLIGAPHYEQTGQGVSVCPLEPRQAPRWQCDVLYGEOGQPKRGGAALT 510

Db 515 NMDGITDFLLVAAPFYHIRGEGRVYVYRV--HEQDAPFSLVYVTSYGLTSRSGFGAMA 573

Qy 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGAVLHFGTSGSGISPSHSQRTAGSKLSPR 564

Db 574 AVGDINQDKFTDVAIGAPLEGFGAGDGASYSVYIYNGHSG-GLHASPQQIRASSVALG 632

Qy 565 LOYFOGSLGGQDLTMDGLVLTIVGAQGHVLLRSQPLRVKAIMFNPREVARNVFECN 624

Db 633 LYVFGMSVSGGLDPFGCDLADITVGSQDVAVYLSRPVVDLTVMSTFTPDALP-----685

Qy 625 DQVVKGEAGEVRVCLHVQKS---TRDLRREGIQSVVTVYDLALDSCGRHSPAVNEFKN 681

Db 686 ---MAFKDNQDVLCFKVDSSAVPEPGLRGMSINFTVDVTV-----TFQ 727

Qy 682 STRRQ-----TQVGLTQTCETIKLQLPNCIEDPVPSPVILRNLP 721

Db 728 KQRLQCADRSQCQSLMKWSGGSSLCHEFGLISTEEL-----CEDDCFSNITIKVSYE 781

```
QY 722 LVGTPLSARGNLRPVLAEDAORLFTALF--PEKNGNDNICODDLSTIFSEMSLDCLVV 779
DB 782 F-OTSEBRRNHPIL--DHYKEPAIFQLPTEKCKKVKVFCIAIQLTALISQOD-LVV 837
QY 780 GGPREFNVTVVRNDEGDSYRTQVTFPPPLDLSYKRVSTLQORQORWRLACESASSTE 839
DB 838 GITKEVTNISTNGSDSYMTNMLNYPNQLQFKI-----QKPLSPDIQCDPKEV- 890
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKXNVTSNNMPTNKTE 899
DB 891 --ASVLWCKLGHPL--KRSSVNVSTWQLESEIIPNRTADITVINSNKSILARETH 947
QY 900 FOLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYVSNLQORSPLISLVPLVP 959
DB 948 --SLOPRHAFVLSR--PSVNYMN--TSQSSDHKEFFNVHGENHFGAVFQLQICVP 1000
QY 960 VRLNQTVHWRQVTFSENLSST-CHTXKELRSHSDFLAELKAPVWVCSIAVCQIQCDD 1018
DB 1001 ETLRDL-----QIIRVHLTKTQART-----ECTSQEPTCGSPVQNV--- 1039
QY 1019 IPFFGIQEEFNATL-----KGNLSFDWYIKTSHNELLI--YSTABIL-----PNDSVFTL 1066
DB 1040 -----BEWHSVICAITSKENVTVAAISMGHTKQLLRDISELQILGEISFNKSLYEG 1092
QY 1067 LFGQGFVRSQVETKEVPP-----EVPNPLIVGSSVGGHLLALITALYKLGFEKQY 1122
DB 1093 LNAE-----NHRKTITVFLKEEKHPSLPLIGSIGGLLVIVIIATLFCGPFPRKY 1146
QY 1123 KDMWSE 1123
DB 1147 KQLNLE 1152

RESULT 15
Q7TQC3
ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.
AC Q7TQC3;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE All integrin.
GN ITGAL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg B.,
RA Gullberg D.;
RT "alpha1 integrin is important for mesenchymal cell function:
RT elimination of alpha1 leads to dwarfism."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RP Johansson M., Popova S.N.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AYL24460; AM62130.3;
XM Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 17.7%; Score 1043.5; DB 11; Length 1188;
Best Local Similarity 28.1%; Pred. No. 1.4e-67;
Matches 349; Conservative 180; Mismatches 495; Indels 219; Gaps 46;

QY 1 FNEIDTENMTFO-ENARGGQSVQ--LOGSR-VVVGAPQEIIVANQORSLYCCDYSGS 56
DB 23 FNMDTNPRVIAGPSAFTGYTQOHDISGKWLTVVGAPMETNGHOKTDGVTKCPVTQGN 82
QY 57 CEPTRL-----QPVVEAVNMSLGLSLAATSPQLLAGPTVHQTCSENTYVKGCLFLF 110
DB 83 CTKLNLRGTVLSNVSEKDNMRGLSLATNPXNSFLACSPLWSHCSSYYTTGMSRV 142
```

```
QY 111 GSNLRQOPQKPEALRGCPQSDSDIAFLIDGSGSIIDHDFRMKEWST-----VMEQLKK- 166
DB 143 NSNFRPSKTVP-ALQRC-QTYXDIVLVDGNSIYP-----WVEQHFILNLIKCF 192
QY 167 ----SKTLPSLMQVSEFRTHFTKFPQNNPNSLKPITQLLG-RHRTATGLKRVRE 221
DB 193 YIFGQIQVGVQVGEDAVHBFHLNDRSVXDVVVEAASHIEQRTGTETRTAFGLEFAESE 252
QY 222 LPNITNGARKNAFKILPLLTDGEKFGDPLGYEDVPELDRGVYRYVIGV-----GDAPR 276
DB 253 AFQ--KGRKGAKKVMIVITDGESHDS-DEKVIQSEKDNVTVYAVAVLGYNNRRGIN 309
QY 277 SEKSRQELNVTASPPRDHVQIINNFEALKTIONQLREKIPIAIEGTQTGSSSSPEHESQ 336
DB 310 PETFLNEIKYTASDPDDKHFFNVTDAAALKDIVDALGDRIEISLEGTNK-NETSGLEMSQ 368
QY 337 EGFSAATISNPLSTVCSYDAGGVFLYTSKEK-----STFINMTRVDSMDNLAIFYA 391
DB 369 TGFSHVVEDGILLGAVGAYDNQAVLKEISAGKVIPIHRESYLKEFPEELKHNHAYLGYT 428
QY 392 AAILLRNV-OSLVGAPRYOHIGLVAMF-RQNTGMWESNANVGTQIGAYFGASLCSVD 449
DB 429 VTSVSSRQGRVYVAGAPRNFHTGKVLFSMHNRSITIHQALRGEQIGSYGSEITVD 488
QY 450 VDSNGSTDLVLIGAPHYYSOTR-GQOVSVCPLPRGQARWQCDVLYGEOQOPWGRFGAA 508
DB 489 VDDRVTDLVLGAPMFPSEGRKGVVYVNL---RQNRFYVNGTLKDSHVSQNAFPGSC 545
QY 509 LTVLGDVNGDKLTVAIGAPCEEDNRGAVYLFHGTSGSGISFSHSQRTAGSKLSPLQVYF 568
DB 546 IASVQDLNODSYNDVWVGAPLEDSHRGAIYIFHGQ-TNILKXPMQRTASELAPGLQHF 604
QY 569 QGSLSGGQDLTMDGVLDTVGAQGHVLLRSQPLRVKAIEMENPREVARNVF--ECNDQ 626
DB 605 GCSHGHQDLNEDGLVDLAVGALGNVAVWAPVQVQINASLHFEPSKI--NIFHKDC--- 659
QY 627 VVKGEAGEVRVCL-----HVQKSTRDLREGQIQSVVTVYDLALDGRPHSR 673
DB 660 ----KNGRDATCLAAFLCFPIFLAPHFQATVVG-----IRYNATMDERYMPR 705
QY 674 AVFNETNS-TRQTVQVGLTQTCTELKLOPNCEBPSVLELNFSLVGTPLSAFGN 732
DB 706 AHLDEGGQFTNRVALLSSGQEHQCRINFHVL-DADYVVPVAFSVESLEDP----- 758
QY 733 LRPLAEDAORLFTALPPEKNGNDNICODDL----- 765
DB 759 NGPMLDNGMPTLRSVFPFNGCNEDEHCVDELVDASDLPTAMEYQCVLGRPAQDCS 818
QY 766 SITFSMELDCLVVGPPREFNVTVVRNDEGDSRTQVTFPPPLDLSYKRVSTLQORQSO 825
DB 819 SVTLSFDTVFIIBSTRRRVAVEATLENRGENAYSAVLNISQSENLP--ASLIQKDDSD 876
QY 826 RSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKX 884
DB 877 NS--IECVNEER-----RLHKVCNVSYFPFRAKAKVAFRLDFEFSKSVFLHLLQHLGA 929
QY 885 NVTSENMMPTNKTEFOLELPVKYAVVMVTVSHGVSTKYLNFTASENTS-----RVM 936
DB 930 GSDSHEQDSTADDNTALLRFLHLYEADVLFTR---SSSLSHFEVKAANSLSYDGIGPFF 986
QY 937 QHQVQVNLGORSPLISLVFL---VPV-----RL-----NOTVIM 968
DB 987 NCVFVKQNLG--FFPIHGVMKMTVPITATRGNNLMLRDFDTGNTSCINWGNSTYR 1044
QY 969 DRPOVTFSENLSSTCHTKERLFSHSD---FLAELKAPVWVCSIAVCQIQCDDIFFGIQ 1025
DB 1045 STET---EEDLS---HAPQRNHSNDVVSIIICNLRAP-----S 1077
QY 1026 BEFNATLGNLSFDWYIKTSHNELLIYSTABI---LFNDSVFTLPGQAFVRSQVETKY 1082
DB 1078 QSTSYLVGNLWLTSEKALKYRSLSKITVNAALQORFHSPIF-----REDDPSRQ 1127
QY 1083 BPFEPVN-----PLPLIVGSSVGGLLLELALITALYKLGPFK 1119
```


Db 1128 VTPEISKOEDQVPIWIIIVGSTLGGELLLALLVIALWKLGFFK 1170

Search completed: June 7, 2004, 17:16:21
Job time : 44.4889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLDTENAWTFQENARGFGQ.....FRRQYKDMVSEGGPPGABPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852	99.5	1153	2 AAW65090	AAW65090 Human Bet
2	5852	99.5	1153	3 AAB07360	AAB07360 Human CD1
3	5852	99.5	1153	5 AAU80252	AAU80252 Human Int
4	5852	99.5	1153	5 ABG61469	ABG61469 Human Bet
5	5852	99.5	1153	5 AAO14428	AAO14428 Integrin
6	5852	99.5	1153	7 ADD25615	ADD25615 Binding d
7	5842	99.3	1153	2 AAR04136	AAR04136 Alpha sub
8	3450	58.6	1163	2 AAR07120	AAR07120 p150.95 a
9	3436	58.4	1163	3 AAW65091	AAW65091 Human Bet
10	3436	58.4	1163	3 AAB07361	AAB07361 Human CD1
11	3436	58.4	1163	5 ABG61470	ABG61470 Human Bet
12	3434	58.4	1163	6 ABU07406	ABU07406 Protein d
13	3401	57.8	1161	2 AAR78166	AAR78166 Human bet
14	3401	57.8	1161	2 AAW23049	AAW23049 Human bet
15	3401	57.8	1161	2 AAW57491	AAW57491 Human bet
16	3401	57.8	1161	2 AAW65089	AAW65089 Human Bet
17	3401	57.8	1161	2 AAW2825	AAW2825 Human alp
18	3401	57.8	1161	2 AAW73342	AAW73342 Human alp
19	3401	57.8	1161	3 AAB07359	AAB07359 Human alp
20	3401	57.8	1161	5 ABG61468	ABG61468 Human Bet
21	3385.5	57.5	1161	2 AAW23064	AAW23064 Human Bet
22	3385.5	57.5	1161	2 AAW65106	AAW65106 Human Bet
23	3385.5	57.5	1161	2 AAW72837	AAW72837 Human alp
24	3385.5	57.5	1161	2 AAW73343	AAW73343 Human alp
25	3385.5	57.5	1161	3 AAB07376	AAB07376 Human alp

26	3385.5	57.5	1161	5 ABG61485	ABG61485 Human Bet
27	3226.5	54.8	1161	2 AAR78169	AAR78169 Rat alpha
28	3224.5	54.8	1161	2 AAW23062	AAW23062 Rat beta
29	3224.5	54.8	1161	2 AAW60004	AAW60004 Rat alpha
30	3224.5	54.8	1161	2 AAW72824	AAW72824 Rat alpha
31	3224.5	54.8	1161	3 AAB07374	AAB07374 Rat alpha
32	3224.5	54.8	1161	5 ABG61483	ABG61483 Rat Beta2
33	3217.5	54.7	1161	2 AAW65104	AAW65104 Rat beta-
34	3217.5	54.7	1161	2 AAW73345	AAW73345 Rat alpha
35	3212	54.6	1161	2 AAW23061	AAW23061 Mouse bet
36	3212	54.6	1161	2 AAW60003	AAW60003 Mouse alp
37	3212	54.6	1161	2 AAW65103	AAW65103 Mouse bet
38	3212	54.6	1161	2 AAW72836	AAW72836 Mouse alp
39	3212	54.6	1161	2 AAW73347	AAW73347 Mouse alp
40	3212	54.6	1161	3 AAB07373	AAB07373 Mouse alp
41	3212	54.6	1161	5 ABG61482	ABG61482 Mouse Bet
42	3208	54.5	1161	2 AAR78168	AAR78168 Mouse alp
43	3203.5	54.4	1151	2 AAW23059	AAW23059 Rat beta
44	3203.5	54.4	1151	2 AAW60001	AAW60001 Rat alpha
45	3203.5	54.4	1151	2 AAW65101	AAW65101 Rat beta-

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX AC AAW65090;

XX XX 28-SEP-1998 (first entry)
XX DT Human Beta-integrin CD11b subunit protein.
XX DE

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; aschma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX OS Homo sapiens.

XX XX US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX XX Screening assay for modulators of integrin binding - using immobilised or
labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX XX Example 5; Fig 1A-D; 106pp; English.

XX XX This sequence represents a human beta-integrin CD11b subunit which is
used to describe a method for identifying compounds that modulate the
interaction of the beta-integrin alpha-d subunit with a binding partner
of alpha-d which involves contacting an alpha-d polypeptide with an alpha
-d binding partner, one of which is immobilised and the other of which is
labelled, in the presence of a test compound, and determining if the
compound affects binding between the alpha-d polypeptide and alpha-d
binding partner, where the alpha-d polypeptide is alpha-d or its fragment
comprising the cytoplasmic, transmembrane or extracellular domain of
alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 CC and rheumatoid arthritis
 XX

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 2; Length 1153;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQSGSVVVGAPQEIIVAAANQRGSVYQCDYSGSCPEI 60
 Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGSVVVGAPQEIIVAAANQRGSVYQCDYSGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATSPQALLACGFTVHTQCSNTYVKGCLFGLGSLNLRQPOK 120
 Db 77 RLQVPEAVNMSLGLSLAATSPQALLACGFTVHTQCSNTYVKGCLFGLGSLNLRQPOK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFFRMKEWSTVMEQLKCKSLFSLMQYSEEF 180
 Db 137 FPEALRGCPQSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKCKSLFSLMQYSEEF 196

Qy 181 RIHPTFKPQNNPNSLIXDITOLLGHTHTATGLRKVRELFNITNGARKNAFKILPLL 240
 Db 197 RIHPTFKPQNNPNSLIXDITOLLGHTHTATGLRKVRELFNITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
 Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPPRDHVFQVN 316

Qy 301 NFEALKTIONQLREKIFAIBGTQTGSSSSPHEHMQSGFSAAITNSGPELLSTVGSYDWAG 360
 Db 317 NFEALKTIONQLREKIFAIBGTQTGSSSSPHEHMQSGFSAAITNSGPELLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRRVQSLVGLAPRYOHIGLVAMFR 420
 Db 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRRVQSLVGLAPRYOHIGLVAMFR 436

Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGGQVSCPL 480
 Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGGQVSCPL 496

Qy 481 PRGORARWOCDAVLYGEGQCPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
 Db 497 PRGORARWOCDAVLYGEGQCPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSORIASKLSPLQVFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 Db 557 HGTSGSGISPSHSORIASKLSPLQVFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLRYKATMEENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRGLREGQIOSVVT 660
 Db 617 PVLRYKATMEENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRGLREGQIOSVVT 676

Qy 661 YDLALDGRPHSRVAFNETKNSRROTQVGLGTQTCETLKLQLENCIEDPVSPIVLRNF 720
 Db 677 YDLALDGRPHSRVAFNETKNSRROTQVGLGTQTCETLKLQLENCIEDPVSPIVLRNF 736

Qy 721 SLVGTPLSAPGNLRFVLAEDAQRFTALFPPEKQKGNNDICQDLSLTFPSMISLDCVLWG 780
 Db 737 SLVGTPLSAPGNLRFVLAEDAQRFTALFPPEKQKGNNDICQDLSLTFPSMISLDCVLWG 796

Qy 781 GPRFENVTVVRNDEGDSYRQVTFPPFDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
 Db 797 GPRFENVTVVRNDEGDSYRQVTFPPFDLSYRKVSTLQNRQSRQSRWLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900
 Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 916

Qy 901 QLELPVKYAVTAVTSHGVSTKYNFTASENTSRVMQHVQVSNLQORSIPISLVLPV 960
 Db 917 QLELPVKYAVTAVTSHGVSTKYNFTASENTSRVMQHVQVSNLQORSIPISLVLPV 976

RESULT 2

AAB07360 standard; protein; 1153 AA.

XX AAB07360;

DT 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

DE Human; macrophage infiltration inhibition; alpha d integrin;
 XX leukocyte integrin; Leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

XX WO200029446-A1.

XX 25-MAY-2000.

XX 16-NOV-1999; 99NO-US027139.

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

XX Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous system
 XX injury sites.

PS Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11b. This sequence
 CC was used in an alignment to identify a novel beta2 integrin alpha
 CC subunit: alpha d (AAA60014 and AAB07359). The present sequence has
 CC approximately 60% identity to the protein sequence of alpha d. The
 CC alpha d gene and protein may be useful in therapy for diseases linked to
 CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 CC (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 CC inhibition of macrophage infiltration at the site of a central nervous
 CC system injury. The monoclonal antibodies can also be used to detect and
 CC diagnose Crohn's disease

SQ Sequence 1153 AA;		Query Match 99.5%; Score 5852; DB 3; Length 1153;		Best Local Similarity 99.2%; Pred. No. 0;		Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;	
QY 1 FNLDTENAMTFOENARGFCQSVVVOIGSSRVVVGAPQEIIVAAANORGLSYQCDYSTGSCPEI 60		DB 17 FNLDTENAMTFOENARGFCQSVVVOIGSSRVVVGAPQEIIVAAANORGLSYQCDYSTGSCPEI 76					
QY 61 RLQVPVEAVNMSGLSLAATTSPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQQQPK 120		DB 77 RLQVPVEAVNMSGLSLAATTSPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQQQPK 136					
QY 121 FFEALRGCPQESDIATFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTILFSLMOYSEEF 180		DB 137 FFEALRGCPQESDIATFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTILFSLMOYSEEF 196					
QY 181 RIHFTFKPQNNPNPSLKIPTQLLGRTHATGLRVKVRRELFNTNGARKNAFLILL 240		DB 197 RIHFTFKPQNNPNPSLKIPTQLLGRTHATGLRVKVRRELFNTNGARKNAFLILL 256					
QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVDAPRSEKSRQELNNTVASKPPRDHVQIN 300		DB 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVDAPRSEKSRQELNNTVASKPPRDHVQIN 316					
QY 301 NFEALXIONQREKIFALIEGTGTGSSSEFHEHMOEGFSAALTNSGPLLSTVGSVDWAG 360		DB 317 NFEALXIONQREKIFALIEGTGTGSSSEFHEHMOEGFSAALTNSGPLLSTVGSVDWAG 376					
QY 361 GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAAILLNRVQSLVLGAPRYOHIGLVAMFR 420		DB 377 GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAAILLNRVQSLVLGAPRYOHIGLVAMFR 436					
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDYDENGSTDLVLIGAPHYRBTQGGQSVCEP 480		DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDYDENGSTDLVLIGAPHYRBTQGGQSVCEP 496					
QY 481 PRGORARWQCDVLYGEOQPMRFGAALTLDVNGDKLTDVAIGAPGEEDNRGAVILF 540		DB 497 PRGORARWQCDVLYGEOQPMRFGAALTLDVNGDKLTDVAIGAPGEEDNRGAVILF 556					
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600		DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 616					
QY 601 PVLRVKAIKEFNPREVARVFCNDQVKGKAGEVRVCLHWKSTRDLRREGQIQSVVT 660		DB 617 PVLRVKAIKEFNPREVARVFCNDQVKGKAGEVRVCLHWKSTRDLRREGQIQSVVT 676					
QY 661 YDLALDSGRPHGRAVFNKNSRTRQTVLGLTQCTETIKLQLEPCIEDPVSPVILRLNF 720		DB 677 YDLALDSGRPHGRAVFNKNSRTRQTVLGLTQCTETIKLQLEPCIEDPVSPVILRLNF 736					
QY 721 SLVGTPLSAFNLRPVLAEDQRLFTALPPEKNCNDNICODDLSITSEFSLDCLVVG 780		DB 737 SLVGTPLSAFNLRPVLAEDQRLFTALPPEKNCNDNICODDLSITSEFSLDCLVVG 796					
QY 781 GPREFNVTYVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWKLACESASTEV 840		DB 797 GPREFNVTYVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWKLACESASTEV 856					
QY 841 SGALKSTCSINHPIIPENSEVTFTFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900		DB 857 SGALKSTCSINHPIIPENSEVTFTFDVDSKASIGNKLLKANVTSENMPRTNKTEF 916					
QY 901 QLELPVKYAVYVWVTSHGVSSTKYNFTASENTSRVWQHQQVQVSNLQQRSLPISLPLVPV 960		DB 917 QLELPVKYAVYVWVTSHGVSSTKYNFTASENTSRVWQHQQVQVSNLQQRSLPISLPLVPV 976					
QY 961 RLNQTVIWRPQVTFPSGNTSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVQRIQCDDIP 1020		DB 977 RLNQTVIWRPQVTFPSGNTSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVQRIQCDDIP 1036					

QY	1021	PFQIQEENATLKGNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPQCGAFVRSQTET	1080
DB	1037	PFQIQEENATLKGNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPQCGAFVRSQTET	1096
QY	1081	KVEPFEVNPPLPLIVGSSVGGLLALITAAALYKLGFFKROYKDMMSGGSPGAEPO	1137
DB	1097	KVEPFEVNPPLPLIVGSSVGGLLALITAAALYKLGFFKROYKDMMSGGSPGAEPO	1153
RESULT 3			
ID	AAU80252	standard; protein; 1153 AA.	
XX	AAU80252;		
DT	15-JUL-2002	(first entry)	
XX	Human integrin 1 alpha-M subunit protein.		
DE	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;		
KW	inflammatory disease; autoimmune disorder; Crohn's disease;		
KW	human immunodeficiency virus; HIV; myocardial infarction;		
KW	Sjorgen's syndrome; rheumatoid arthritis.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 499..500	/note= "Encoded by GGG CAG AGG"	
XX	WC020218583-A2.		
PD	07-MAR-2002.		
XX	31-AUG-2001; 2001WO-US027227.		
XX	01-SEP-2000; 2000US-0229700P.		
PA	(BL00-) CENT BLOOD RES INC.		
PI	Springer TA, Shimoaka M, Lu C;		
XX	WPI; 2002-382964/41.		
DR	N-PSDB; ABX50046.		
XX	Modified integrin-I or integrin I-like domain polypeptide useful as an immunogen to produce antibodies specific to polypeptide, comprises a disulfide bond such that polypeptide is stabilized in a desired conformation.		
PS	Disclosure; Page 109-112; 112pp; English.		
CC	This invention relates to a modified integrin-I or integrin I-like domain polypeptide comprising at least one disulfide bond so that the domain is stabilised in a desired conformation. The polypeptide of the invention may have antiinflammatory or immunosuppressive activities. The polypeptides of the invention have an open conformation and are useful as immunogens to produce antibodies that selectively bind to integrin I-domain; and for identifying a modulator of integrin activity, or of interaction of an integrin and a cognate ligand. The polypeptide of the invention, or antibodies (preferably anti-LFA-1 antibody) is useful for treating or preventing an integrin mediated disorder which is an inflammatory or autoimmune disorder in a subject and for inhibiting the binding of an integrin to a cognate ligand such as Crohn's disease, nephritis; human immunodeficiency virus (HIV), myocardial infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic composition comprising the peptide of the invention is useful for treating an integrin mediated disorder in a subject. The polypeptides and/or active or antigenic fragments are useful as reagents for diagnosis of integrin-mediated disorders. The present sequence represents the human integrin-1 alpha-M protein subunit used to generate the mutant polypeptides of the invention		

SQ		Sequence 1153 AA;	
Query Match		99.5%; Score 5852; DB 5; Length 1153;	
Best Local Similarity		99.2%; Pred. No. 0;	
Matches 1128; Conservative		7; Mismatches 2; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTFOENARGPQSQSVVQLQGSRVVVGAPQEIIVAAANQORSLVQCDYSTGSCBPI	60
DB	17	FNLDTENAMTFOENARGPQSQSVVQLQGSRVVVGAPQEIIVAAANQORSLVQCDYSTGSCBPI	76
QY	61	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHOTCSSENTYVKGCLFGLGSLNLRQOQPK	120
DB	77	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHOTCSSENTYVKGCLFGLGSLNLRQOQPK	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKRWVSTVMEQLKSKTFLSLMQYSEEF	180
DB	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKRWVSTVMEQLKSKTFLSLMQYSEEF	196
QY	181	RIHFTFKFQNNPNRSLKPIITQLLGRTHRTATGLRVKVRRELFNTNGARKNAFKILPIL	240
DB	197	RIHFTFKFQNNPNRSLKPIITQLLGRTHRTATGLRVKVRRELFNTNGARKNAFKILVVI	256
QY	241	TGEXFGDPLGYEDVPELDRGVIRYVIGVGDADFSEKSRQELNTVASKPPDRDHVFQIN	300
DB	257	TGEXFGDPLGYEDVPELDRGVIRYVIGVGDADFSEKSRQELNTVASKPPDRDHVFQIN	316
QY	301	NFEALXTIONLREKIPAFIEGTQTGSSSFHEMSQEGFSAITNSGPLSTVGSYDWAG	360
DB	317	NFEALXTIONLREKIPAFIEGTQTGSSSFHEMSQEGFSAITNSGPLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSPINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPRYQHIGLVAMFR	420
DB	377	GVFLYTSKEKSPINNTRVDSMDNDAYLGYAAAILRNVRQSLVGLGAPRYQHIGLVAMFR	436
QY	421	QNTGMFESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQVSVCP	480
DB	437	QNTGMFESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQVSVCP	496
QY	481	PGQARWQCDVAVLGEQGPGRFCAALTVLGDVNGDKLTVAICAPGEENRGAVYLP	540
DB	497	PGQARWQCDVAVLGEQGPGRFCAALTVLGDVNGDKLTVAICAPGEENRGAVYLP	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ	616
QY	601	PVLVRKALIMEFPRVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQIQSVVT	660
DB	617	PVLVRKALIMEFPRVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQIQSVVT	676
QY	661	YDLALDSGRPHGRAVFNKSTRQTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF	720
DB	677	YDLALDSGRPHGRAVFNKSTRQTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF	736
QY	721	SLVGTPLSAFGMLRPLVLAEDAQLFTALPPFKNGNDNICODDLSITPFSNLSCLVVG	780
DB	737	SLVGTPLSAFGMLRPLVLAEDAQLFTALPPFKNGNDNICODDLSITPFSNLSCLVVG	796
QY	781	GPREFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSNRLACESASSTEV	840
DB	797	GPREFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSNRLACESASSTEV	856
QY	841	SGALKSTCSGNIHPIPPENSEVTFTNTPDVSASLGNKLLKANVTSENPNRKNTEF	900
DB	857	SGALKSTCSGNIHPIPPENSEVTFTNTPDVSASLGNKLLKANVTSENPNRKNTEF	916
QY	901	QLELPVKYAVYVMTSHGVSYKLYNFTASENTSRVMOHOYQVSNLQORSLPLSLVFLVPV	960
DB	917	QLELPVKYAVYVMTSHGVSYKLYNFTASENTSRVMOHOYQVSNLQORSLPLSLVFLVPV	976
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVQCRIQCDIP	1020
DB	977	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVQCRIQCDIP	1036
QY	1021	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQCGAFVRSQTEP	1080
DB	1037	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQCGAFVRSQTEP	1096
QY	1081	KVEPFEVNPFLVIGSVGGILLALITAAALYKLGFFKROYKDMSEGGPPGABPQ	1137
DB	1097	KVEPFEVNPFLVIGSVGGILLALITAAALYKLGFFKROYKDMSEGGPPGABPQ	1153
RESULT 5			
AAO14428		standard; protein; 1153 AA.	
XX		AAO14428;	
XX		03-MAY-2002 (first entry)	
DT		Integrin Mac-1 alpha subunit.	
DE		Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
XX		open conformation; integrin related inflammatory disorder;	
KW		integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW		reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW		viral infection; cancer; gene therapy; vaccine;	
KW		bioactive agent screening.	
XX		Unidentified.	
OS		WO200204521-A2.	
XX		17-JAN-2002.	
PD		09-JUL-2001; 2001WO-US021805.	
XX		07-JUL-2000; 2000US-0216600P.	
XX		(CALY) CALIFORNIA INST OF TECHNOLOGY.	
PA		(BLOO-) CENT BLOOD RES.	
XX		Springer T;	
XX		WPI; 2002-148167/19.	
DR		New integrin I domain protein having alteration in at least 2	
PT		noncontiguous regions and exits in an open conformation, useful for	
PT		treating, preventing or suppressing inflammatory or immunological	
PT		disorders.	
XX		Example 1; Fig 1P; 90pp; English.	
PS		The invention comprises structurally biased variant integrin inserted (I)	
XX		domain proteins, wherein the alterations to the protein occur in at least	
CC		two noncontiguous regions. Specifically the variant integrin I domain	
CC		proteins are structurally biased to exist in the open conformation.	
CC		thereby altering the binding ability of the protein. The invention also	
CC		comprises nucleic acids encoding the variant integrin I domain proteins.	
CC		The integrin I domain proteins and nucleic acids are useful for treating,	
CC		preventing or suppressing integrin related inflammatory and immunological	
CC		disorders (e.g. rheumatoid arthritis). The variant integrin I domain	
CC		proteins and nucleic acids can also be used for treating: ischaemia/	
CC		reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral	
CC		infection; and cancer. The variant integrin I domain nucleic acids and	
CC		proteins may be used in gene therapy, as vaccines and to screen for	
CC		bioactive agents. The present amino acid sequence represents the Mac-1	
CC		alpha subunit of integrin	
XX		Sequence 1153 AA;	
SQ		Query Match 99.5%; Score 5852; DB 5; Length 1153;	
		Best Local Similarity 99.2%; Pred. No. 0;	
		Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;	

QY 1 FNLDTENAMTFOENARGFGQSVVLQGSRRVVGAPQEI VAAANORGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVLQGSRRVVGAPQEI VAAANORGLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQILLACGPTVHTQCSNTYVYKGLCFPLPSSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQILLACGPTVHTQCSNTYVYKGLCFPLPSSNLRQPOK 136
QY 121 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDPRRMEKVVSTVMEQLKSKSTLPSLMQYSEBF 180
DB 137 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDPRRMEKVVSTVMEQLKSKSTLPSLMQYSEBF 196
QY 181 RIHFTKFEQNNPNPSLKIPIITQLLGRTHATGLRKKVRELFNITNGARKNAKILFLL 240
DB 197 RIHFTKFEQNNPNPSLKIPIITQLLGRTHATGLRKKVRELFNITNGARKNAKILFLL 256
QY 241 TDEKEGDPGLYEDVPIPEADREGVIRYVIGVGDFAFSEKSRQELNTIASKPPROHVQVFN 300
DB 257 TDEKEGDPGLYEDVPIPEADREGVIRYVIGVGDFAFSEKSRQELNTIASKPPROHVQVFN 316
QY 301 NFALKTIONOLREKIPAIETGOTGSSSPHEHMSQEGFSAITSNGLPSTVGSYDWAQ 360
DB 317 NFALKTIONOLREKIPAIETGOTGSSSPHEHMSQEGFSAITSNGLPSTVGSYDWAQ 376
QY 361 GVFLYTSKESKSTINTRVDSMDNDAYLVYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTINTRVDSMDNDAYLVYAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVGTQIGYFAGSLCSDVDNSGSDTLVLIGAPHYYEOTRGQSVSCPL 480
DB 437 QNTGMWESNANVGTQIGYFAGSLCSDVDNSGSDTLVLIGAPHYYEOTRGQSVSCPL 496
QY 481 PRGORARWQCDVLYGEOGPWREGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLP 540
DB 497 PRGORARWQCDVLYGEOGPWREGALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLP 556
QY 541 HGTSGSISPSHSORAGSKSLSPRLAYFGOSLGGQDLTMDGLDVLTVGAQGHVLLLRQ 600
DB 557 HGTSGSISPSHSORAGSKSLSPRLAYFGOSLGGQDLTMDGLDVLTVGAQGHVLLLRQ 616
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKKEAGEVRVCLHVKSTRDLRRCQIOSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFCNDQVVKKEAGEVRVCLHVKSTRDLRRCQIOSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGRNRPVLARDAQRLFTALPPFKNCGNDNICDDLSITFSFXGLDCLVVG 780
DB 737 SLVGTPLSAFGRNRPVLARDAQRLFTALPPFKNCGNDNICDDLSITFSFXGLDCLVVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLLACESASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTKTEP 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTKTEP 916
QY 901 QLELPVYAVYVTVSHGVSTKYNLFTASNTSRVMQHYOVSNLQORSPLISLVLVPV 960
DB 917 QLELPVYAVYVTVSHGVSTKYNLFTASNTSRVMQHYOVSNLQORSPLISLVLVPV 976
QY 961 RENQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAQVWNSIAVCORIQCIP 1020
DB 977 RENQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAQVWNSIAVCORIQCIP 1036
QY 1021 PRGOBENATLKNLSFDVWIKTSHHLLIVSTAEILLFNDVSFTLLPGQAFVRSOTET 1080
DB 1037 PRGOBENATLKNLSFDVWIKTSHHLLIVSTAEILLFNDVSFTLLPGQAFVRSOTET 1096
QY 1081 KVEPFEVFNPLTVGVSSVGLLALLALITAAALYKLGFFKQYKDMMSGGPPGAPQ 1137

DB 1097 KVEPFEVFNPLTVGVSSVGLLALLALITAAALYKLGFFKQYKDMMSGGPPGAPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX AC ADD25615;
XX DT 15-JAN-2004 (first entry)
XX DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX KW Binding domain; immunoglobulin; fusion protein; cytostatic;
XX KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
XX KW neuroprotective; hinge region; immunoglobulin heavy chain;
XX KW CH2 constant region; CH3 constant region; IgG1;
XX KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
XX KW rheumatoid arthritis; B-cell disorder; melanoma; carcinoma; sarcoma;
XX KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX OS Unidentified.
XX PN US2003118592-A1.
XX PD 26-JUN-2003.
XX PF 25-JUL-2002; 2002US-00207655.
XX PR 17-JAN-2001; 2001US-0367358P.
XX PR 17-JAN-2002; 2002US-00053530.
XX PR 03-JUN-2002; 2002US-0385691P.
XX PA (GENE-) GENE-CRAFT INC.
XX PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX DR WPI; 2003-801317/75.
XX PT New binding domain-immunoglobulin fusion protein, useful for treating a
XX PT subject having or suspected of having a malignant condition or a B-cell
XX PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX PS Disclosure; SEQ ID NO 176; 157pp; English.
XX CC The invention relates to a binding domain-immunoglobulin fusion protein
XX CC comprising a binding domain polypeptide that is fused to an
XX CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
XX CC CH2 constant region polypeptide that is fused to the hinge region
XX CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
XX CC polypeptide that is fused to the CH2 constant region polypeptide. The
XX CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
XX CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
XX CC region polypeptide; derived from (a) having 3 or more cysteine residues;
XX CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
XX CC contains 2 cysteine residues, where the first cysteine is not mutated; a
XX CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
XX CC (a) having 3 or more cysteine residues, where the mutated human IgG1
XX CC immunoglobulin hinge region polypeptide contains no more than one
XX CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
XX CC polypeptide, derived from (a) having 3 or more cysteine residues; where
XX CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
XX CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
XX CC capable of at least one immunological activity comprising antibody
XX CC binding domain polypeptide is capable of specifically binding to an
XX CC antigen. Also included are an isolated polynucleotide encoding the
XX CC binding domain-immunoglobulin fusion protein, a recombinant expression
XX CC construct comprising the polynucleotide operably linked to a promoter,
XX CC a host cell transfected or transfected with a recombinant expression
XX CC construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin
CC fusion protein or polynucleotide and a carrier and treating a subject
CC having or suspected of having a malignant condition or a B-cell disorder.
CC The binding domain-immunoglobulin fusion protein is useful for treating a
CC subject having or suspected of having a malignant condition or a B-cell
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC sclerosis or autoimmune disease. The present sequence is a binding domain
CC -immunoglobulin fusion protein-associated protein sequence. Note: The
CC sequence data for this patent formed part of the printed specification
CC and is also available in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC identified the sequences in the printed specification by their SEQ ID
CC number therefore none of the sequences can be explicitly identified.
XX
SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 7; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFENARGFGQSVVQLQGSVVVGAQOEIVAAHQRSGLVQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFENARGFGQSVVQLQGSVVVGAQOEIVAAHQRSGLVQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVXGLCFGLFGLNLRQOQPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVXGLCFGLFGLNLRQOQPK 136
QY 121 PFEALRGCEQSDTAFLDGGSGIIPHDPRFKWVSTVMEOLKSKTLPFLMLOYSEBF 180
DB 137 PFEALRGCEQSDTAFLDGGSGIIPHDPRFKWVSTVMEOLKSKTLPFLMLOYSEBF 196
QY 181 RIHFTFKFQNNPNRSLKIPITQLGRTHATGLRVKRVRELNFNTNGARKNAKILFIL 240
DB 197 RIHFTFKFQNNPNRSLKIPITQLGRTHATGLRVKRVRELNFNTNGARKNAKILVVI 256
QY 241 TGEKFGDPLGYEDVITPELDRGVIRYVGVGADPRSEKSRQELNVTASKPPRDRHVFQIN 300
DB 257 TGEKFGDPLGYEDVITPEADRGVIRYVGVGADPRSEKSRQELNVTASKPPRDRHVFQIN 316
QY 301 NFEALKTIONQLREKIFALEGTQTGSSSFHEMSQEGFSAITNGPLLSVTGSDYDAG 360
DB 317 NFEALKTIONQLREKIFALEGTQTGSSSFHEMSQEGFSAITNGPLLSVTGSDYDAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILLRNRVQSLVIGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILLRNRVQSLVIGAPRYQHIGLVAMPR 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCEL 480
DB 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCEL 496
QY 481 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVYLF 540
DB 497 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQHVLLRQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQHVLLRQ 616
QY 601 PVLRVKAIMFNPREVARNVFCNDQWVGKXAGEVRVCLHWKSTRDRLRGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFCNDQWVGKXAGEVRVCLHWKSTRDRLRGQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNETHKSTRTOVNLGLTQTCETIKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSAVFNETHKSTRTOVNLGLTQTCETIKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAPGNLRPVLAEQAQLFTALPPFENKCGNDNICODIASITPSFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEQAQLFTALPPFENKCGNDNICODIASITPSFMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHIFPENSEVFNITFDVDSKASLGKLLIKANVTSENNKMPSTNKTET 900
DB 857 SGALKSTSCSINHIFPENSEVFNITFDVDSKASLGKLLIKANVTSENNKMPSTNKTET 916
QY 901 QLELPVKYAVVWVTVSHGVSTKYLNFPTASENTRSVMOHQYQVSNLQORSPLISLVFLVPV 960
DB 917 QLELPVKYAVVWVTVSHGVSTKYLNFPTASENTRSVMOHQYQVSNLQORSPLISLVFLVPV 976
QY 961 RLNTQVINDRPQVTFSENLSSTCHTKRSLPSHSDFLAELRKAPVWVNCISIAVCRIQCDIP 1020
DB 977 RLNTQVINDRPQVTFSENLSSTCHTKRSLPSHSDFLAELRKAPVWVNCISIAVCRIQCDIP 1036
QY 1021 FSGIOBEFNATLKGNSLDFWDYIKTSHNHLIVSTABILLFNDVSFTLLPGQGFVRSQDET 1080
DB 1037 FSGIOBEFNATLKGNSLDFWDYIKTSHNHLIVSTABILLFNDVSFTLLPGQGFVRSQDET 1096
QY 1081 KVPEFVENPLPLIVGSSVGGILLALITAAALYKLGFPFKQYKDMSEGPPGABPQ 1137
DB 1097 KVPEFVENPLPLIVGSSVGGILLALITAAALYKLGFPFKQYKDMSEGPPGABPQ 1153
RESULT 7
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
AC AAR04136;
XX
DT 25-MAR-2003 (revised)
DT 07-SEP-1990 (first entry)
XX
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW non-specific defence system; integrin gene superfamily.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FH Region 1..16
FT /label= signal peptide
FT Modified-site 86..88
FT /label= putative N-glycosylation site
FT Modified-site 240..242
FT /label= putative N-glycosylation site
FT Modified-site 391..393
FT /label= putative N-glycosylation site
FT Modified-site 469..471
FT /label= putative N-glycosylation site
FT Modified-site 693..695
FT /label= putative N-glycosylation site
FT Modified-site 697..699
FT /label= putative N-glycosylation site
FT Modified-site 735..737
FT /label= putative N-glycosylation site
FT Modified-site 802..804
FT /label= putative N-glycosylation site
FT Modified-site 881..883
FT /label= putative N-glycosylation site
FT Modified-site 901..903
FT /label= putative N-glycosylation site
FT Modified-site 912..914
FT /label= putative N-glycosylation site
FT Modified-site 941..943
FT /label= putative N-glycosylation site
FT Modified-site 947..949
FT /label= putative N-glycosylation site
FT Modified-site 979..981
FT /label= putative N-glycosylation site
FT Modified-site 994..996

FT Modified-site /label= putative N-glycosylation site
FT 1022..1024
FT /label= putative N-glycosylation site
FT Modified-site 1045..1047
FT /label= putative N-glycosylation site
FT Modified-site 1051..1053
FT /label= putative N-glycosylation site
FT Modified-site 1076..1078
FT /label= putative N-glycosylation site
FT 1106..1134
FT /label= putative transmembrane region
XX
PN SP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-00115159.
XX
XX 23-AUG-1988; 88US-00235353.
PR 09-MAR-1989; 89US-00321239.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Springer TA, Corbi A;
XX
XX WPI; 1990-125938/17.
DR N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
FT inflammation and viral infections, and in diagnosis.
FT
XX Disclosure; Page ?; -pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC recognition of and migration to sites of inflammation. It also attaches
CC to cellular substrates as part of this function making it useful in
CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC -MAR-2003 to correct PA field.)
XX
XX Sequence 1153 AA;

Query Match 99.3%; Score 5842; DB 2; Length 1153;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQBIIVANORGSLSYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLGSRVVGAPQBIIVANORGSLSYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMISGLSLAATTPPQLACGPTVHQTCSNTYVKGLCFGLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMISGLSLAATTPPQLACGPTVHQTCSNTYVKGLCFGLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDIAFLIDGSGIIPHDPRMKENVSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 PPEALRGCPQSDIAFLIDGSGIIPHDPRMKENVSTVMEQLKSKTLFSLMOYSEEF 196
Qy 181 RIHFTFEFQNNPRSLIKPIITOLLGRTHATGLRVKWRVRELFTNGARKNAFKILPL 240
Db 197 RIHFTFEFQNNPRSLIKPIITOLLGRTHATGLRVKWRVRELFTNGARKNAFKILV 256
Qy 241 TDGKFKGDLGYEDVIBELDREGVIRVVGDAFRSEKSRQELNTVASKPRDHVFPQIN 300
Db 257 TDGKFKGDLGYEDVIBELDREGVIRVVGDAFRSEKSRQELNTVASKPRDHVFPQV 316
Qy 301 NFEALKTIONLRREKIPIAETGTQGTSSSSPHEMSQGFSAIITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONLRREKIPIAETGTQGTSSSSPHEMSQGFSAIITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSDMDAYGLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSDMDAYGLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTRGGQSVCP 496
Qy 481 PRGQARWQCDVAVLYGEOGQPWGRFGAALTVLGDVNGKLTDLVAIGAPGEEHNRGAVYLF 540
Db 497 PRGQARWQCDVAVLYGEOGQPWGRFGAALTVLGDVNGKLTDLVAIGAPGEEHNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKLSPLRQYFGQSLSGGQDLTMDGLVLTIVGAQGHVILLARSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLRQYFGQSLSGGQDLTMDGLVLTIVGAQGHVILLARSQ 616
Qy 601 PVLRVKAIKEFNPREVARNVFECDNDQVVKGEAGVRVCLHVOKSTRDLRREGQIOSVVT 660
Db 617 PVLRVKAIKEFNPREVARNVFECDNDQVVKGEAGVRVCLHVOKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGUTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGUTQTCETLKLQLPNCIEDPVPVILRLNF 736
Qy 721 SLVGTPLSAFNLRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITPSPMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITPSPMSLDCLVVG 796
Qy 781 GPREFNVTVVENDGEDSVRTQVTFPEPDLISYRVKUSTLONORSORSNWLACESASSTEV 840
Db 797 GPREFNVTVVENDGEDSVRTQVTFPEPDLISYRVKUSTLONORSORSNWLACESASSTEV 856
Qy 841 SCALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKAKVTSNNMPRTNKTEF 900
Db 857 SCALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKAKVTSNNMPRTNKTEF 916
Qy 901 QLELPVKVAVVTVSHGVSTKYLNFASNTSRVMQHOYQVNSLNGORSLSPLSLVFLVPV 960
Db 917 QLELPVKVAVVTVSHGVSTKYLNFASNTSRVMQHOYQVNSLNGORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSCTHKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSCTHKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1036
Qy 1021 PFGIOEFPNATLKGMLSPDWYIKTSHNHLIIIVSTAILFENDSVFTLLPGCGAFVRSQTE 1080
Db 1037 PFGIOEFPNATLKGMLSPDWYIKTSHNHLIIIVSTAILFENDSVFTLLPGCGAFVRSQTE 1096
Qy 1081 KVEPPVPVPLPLIVGSSVGGILLALITAAALYKLGFFKRVQKMDMMSEGGPPCAEPQ 1137
Db 1097 KVEPPVPVPLPLIVGSSVGGILLALITAAALYKLGFFKRVQKMDMMSEGGPPCAEPQ 1153
RESULT 8
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
XX 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
XX p150.95 alpha subunit encoded by clone lambdaX47.
XX
XX p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
XX rhinovirus.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX 1..19
XX Region /label= signal peptide
XX 20..44
XX Region /label= N-terminus
XX 61..63
XX Modified-site /label= glycosylation site

PD 25-MAY-2000.
 XX 16-NOV-1999; 99WO-US027139.
 XX 16-NOV-1998; 98US-00193043.
 PR 08-JUL-1999; 99US-00350259.
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 XX WPI; 2000-387751/33.
 DR Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous system
 PT injury sites.
 XX Example 5; Fig 1; 270pp; English.
 XX Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This sequence
 CC was used in an alignment to identify a novel beta2 integrin alpha
 CC subunit: alpha d (AA60014 and AA607359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha d. The
 CC Alpha d gene and protein may be useful in therapy for diseases linked to
 CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 CC inhibition of macrophage infiltration at the site of a central nervous
 CC system injury. The monoclonal antibodies can also be used to detect and
 CC diagnose Crohn's disease
 XX Sequence 1163 AA;
 SQ
 Query Match 58.4%; Score 3436; DB 3; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 1.2e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;
 1 FNLDZENAMTFQENARGFQSGSVVQGGSRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
 20 FNLDTEELTAERVDAGFQSGSVVQVANSVWVGAPQKIIAANQIGLYQCYSTGACBPI 79
 61 RLQVPEAVNNSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQOQK 120
 80 GLQVPEAVNNSLGLSLASTTSPQLACGPTVHHECGENMYLTGLCLLQPT.-QLTQR 137
 121 FPEALRGCCQEBSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTLPISLMQYSEEF 180
 138 LPVSRQECPRQDQIVFLIDGSGSISSRNFAFMVAVISQFORPSTQFSLMQFSNK 197
 181 RHFTFKFQNNPNRSLKPTQLLGRTHATGRLKVVVRELFINITNGARKNAFKILFL 240
 198 QTHFTFEEFRRTSNPLSLASVHLQSGFTYTATQNVVHRLPHASYGARRDAIKILVI 257
 241 TDCEKFGDPLGYEDVPELDREGVIRYVGVGDAPESKSRQELNATVASKPPDRHVFOIN 300
 258 TDCKEGSDLDYVDFPMADAGIIRYALGVGLAFQNRNSWELNDIAKSPQSEIFKVE 317
 301 NFEALXTIQNLRKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLSTVGSYDAG 360
 318 DEFALDKIQNLRKIFAIEGTETSSSFELEMAQEGFSAVFTPDGVLGAVGQFTWSG 377
 361 GYFLYTSKESFTINRVDNDVILCYAAAIILRNQVSLVLAGPRYQHILVAMFR 420
 378 GAFLYPPNMSPTFINNSQENVDNRDSYGLSYTELALWKGVSVLGAPRYQHIGKAVIFI 437
 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTREGQVSCVCLP 480

Db 438 QVSRQMRKAEVIGTQIGSYFGASLCSDVDVDTGSGTDLVLIGAPHYYEQTREGQVSCVCLP 497
 QY 481 PRGORARWOCDAVLYGEOGQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVYLF 540
 Db 498 PRGWR-RWWDVAVLYGEOGHPGRFGAALTVLGDVNGDKLTDVVI GAPGEENRGAVYLF 556
 QY 541 HGTSGSISPSHSORIASKLSPLRQYFGQSLSGGQDLTMDGLVDLTIVGAGCHVLLRSQ 600
 Db 557 HGVLPSPISPSHSORIASQSLSRLOQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
 QY 601 PVLRYKALMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVOKSTDRRLRGOIQSVVT 660
 Db 617 PVLWVGSMQIPAEIPRPAFECNEQVVSQTLVQSNICLYIDKRSKNLGRDLQSSVT 676
 QY 661 YDLALDSGRPHSRAVENETNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
 Db 677 LDALAPGRLSRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRLNF 736
 QY 721 SLVCTPLSACNLRPVLAEQAORLPTALFPPEKNCNDNICODDLISITFSFMSLDCLVVG 780
 Db 737 TLVGKPLLAFLANLPMALAAQRYFTASLPFKNCAGADHICODNLGISFSFGLKSLIVG 796
 QY 781 GPREFNVTVTRNDGDSYRTQVTFPPFLDLISYKVTQLQNRQSQRWRLACESASSTEV 840
 Db 797 SNLENAEAVWVNDGDSYGTITTFSHPAGLSYRYVAEQKQGLRSILHLC--CSAPVG 854
 QY 841 SCALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLIKANVTSENNMPRTNTEF 900
 Db 855 SQGTWSTSCRINHILIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSENINPTSKTIF 914
 QY 901 QLELPVAVVAVVWVTSHGVS TKYLNFTAS-ENTSRVQHOYQVSNLQORSIPISLVLVP 959
 Db 915 QLELPVAVVAVVVSSEHQPTKYLNFSSESEKSHVAMHRVQVNVNLQORDLPVSNFVWP 974
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTVERLPSSHDELAEALRKAVNCSIAVCORICDI 1019
 Db 975 VELNQEAVMMDVEVSHFPQNPISRCSEKIAAPPASDFLAHQKPNVLDCSIAGCLFRCDV 1034
 QY 1020 PFGIQEENATLKGSLSPDWYIKTSHNHLIVSTABILPNDSTVFTLLPGCGAFVRSQTE 1079
 Db 1035 PSFSVQELDFTLKGNLSFGWVRQILQKQVSVVSAIIFDTSVYVQLPGQEFWEAQTI 1094
 QY 1080 TKVPPFVENPLPLIVGSSVGLLLALITLALITLALITLALITLALITLALITLALITL 1128
 Db 1095 TVLEKYKVHNPILIVGSSIGGLLLALITLALITLALITLALITLALITLALITLALIT 1143
 RESULT 11
 ABG61470 standard; protein; 1163 AA.
 XX AC ABG61470;
 XX DT 27-AUG-2002 (first entry)
 XX Human Beta2 integrin alphaCD11c subunit.
 DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; IAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-2; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX WO2002030980-A2.
 XX 18-APR-2002.
 XX

FT Region /note= "extracellular domain"
 FT 150..352
 FT /note= "this region is homologous to the insert common to
 FT Chila,b,c and may be a site for interaction with ICAM
 FT family proteins"
 FT 465..474
 FT /note= "putative cation binding site"
 FT 518..527
 FT /note= "putative cation binding site"
 FT 592..600
 FT /note= "putative cation binding site"
 FT 1109..1128
 FT /note= "transmembrane region"
 FT 1129..1161
 FT /note= "cytoplasmic domain"
 FT XX
 PN W09517412-A1.
 XX 29-JUN-1995.
 XX 21-DEC-1994; 94WO-US014832.
 XX 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1995-240603/31.
 DR N-PSDB; AAQ91712.
 DR Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 XX Claim 7; Page 82-87; 172pp; English.
 CC A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -TM1 was used to screen a human spleen cDNA library to identify clone
 CC 1A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO cells
 XX Sequence 1161 AA;
 SQ Query Match 57.8%; Score 3401; DB 2; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 1e-275;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFQNAFGQSVVQVQGSRLVVGAPLEVVAAQTGRLYDCAATGMCQPI 60
 DB 17 FNLDVEEPTIFQDAGFGQSVVQVQGSRLVVGAPLEVVAAQTGRLYDCAATGMCQPI 76
 QY 61 RLOVPVEAVNMISGLIAATTSPQACGPTVHCQSCENTYVKGCLFLFGNLQPOPK 120
 DB 77 PLHIREAVNMISGLIAATTSPQACGPTVHCQSCENTYVKGCLFLFGNLQPOPK 135
 QY 121 FPEALRGCPQEDSDIAFLDGSGSIIPHPFRNKEMVSTVMEQLKSKTFLSVMQSEEF 180
 DB 136 VPDATPECPHOEMDIVFLDGSGSIQDNDNFQMGFQVQVQMGQFEGDTFLFALMOYSNLL 195
 QY 181 RHFTKEPQNNPNSLKIPIQALGRHTATGLKRVVRELFINITNGARKNAKPIFLFL 240
 DB 196 KHFTPTQFSTSSQSLSDPIVOLKGLIFTATGILTVTQLPHHKNGARKSAKILLIYI 255
 QY 241 TDGEKFGDPLGYEDVPIELDREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFQIN 300
 DB 256 TDGQKYKDPLEYSVPIQAEKAGIIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFQIN 315
 QY 301 NFPAALQIQLOKQEKIYAVEGTQSRASSSSFOHEMSQEGFSTALTMQDPLGAVGSFWSG 360
 DB 316 NFPAALQIQLOKQEKIYAVEGTQSRASSSSFOHEMSQEGFSTALTMQDPLGAVGSFWSG 375

QY 361 GVFLYTSKESKSTFNMTVRVSDMDNDAYLYGAAAILLNRVQSLVGLGAPRYCHIGLVAVER 420
 DB 376 GAFLYPPMSPFTFNMSQENVMDRDSYLGYSYSLALWKGVQNLVGLAPRQHTKAVIPT 435
 QY 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHYEQTRGGQVSCPL 480
 DB 436 QVSRQWRKKAARVTCQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYHYEQTRGGQVSCPL 495
 QY 481 PRQARWQCDVLYGEGQPGWGFNAALTVLDVNGCKLTDVAIGAPGEDNRGAVYLF 540
 DB 496 PRQORVQMCDAVLRGEGQHPWGFNAALTVLDVNGCKLTDVAIGAPGEDNRGAVYLF 555
 QY 541 HGTSGGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLRLRSQ 600
 DB 556 HGASEGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLRLRSQ 615
 QY 601 PVLRYKAIMFNPREVARNVFCNDQVVKGEKAGEVRVCLHVQKSTDRLEBQIOSVVT 660
 DB 616 PVLKVGVMRFPSPVEVAKAVYRWEERKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRAPVNETKNSRTOVQLGLTCTETLKLQFPNCIEDPVSPIVLRINF 720
 DB 674 FDLALDPGLRLTSRAIFNETKNPTLTKRKTGLGTHCHETLKLALLPDCVEDVSPILHLNF 733
 QY 721 SLVGTPLSAFENLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITTFSPMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFKKNCNDNICODDLSITTFSPMSLDCLVVG 793
 QY 781 GPREFNTVTVRNDGEDSYRTOVTFPPPLDLSYKVKSTLQNRQSRQSRWRLACESASSTEV 840
 DB 794 SSLENLVIVVNNAGEDSYRTOVTFPPPLDLSYKVKSTLQNRQSRQSRWRLACESASSTEV 852
 QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENKMPKNTKTEP 900
 DB 853 EG-LRSSRCSVNHPIFHEGSGNGTPIVTFDVSKATLGDRLMRASSENKTKASSSKATF 911
 QY 901 QLELPVXVAVMVTSHGVSTKYNLF--TASENTSRVQHOYOVSNLQSRSLPISLVFLVP 959
 DB 912 QLELPVXVAVMVTSHGVSTKYNLF--TASENTSRVQHOYOVSNLQSRSLPISLVFLVP 971
 QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTRELPSHSDPLAELKAPVAVNCISIANVQRIQCDI 1019
 DB 972 VLLNGVAVWDVMEAPSQSL--PCVSRKPPQHSDFLTQTSRSPMLDCSIADCLQFRCDV 1029
 QY 1020 PFGIQBEPNATLKGNSLFDWYIKTSHNLLIVSTABILENDSVPTLLPGOGAFVRSQTE 1079
 DB 1030 PFSVQBELDFTLKGNSLFGWVRETQKVLVSVABEITDTSVYSQLPQGGAPFRAQNE 1089
 QY 1080 TKVPEFVEPNPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSE 1128
 DB 1090 MVLSEDEVYNAIPIIMGSSVCGALLLALITATLYKLGFFKSHYKEMLED 1138
 RESULT 14
 AAQ23049
 ID AAQ23049 standard; protein; 1161 AA.
 XX AAQ23049;
 AC AAQ23049;
 DT 24-FEB-1998 (first entry)
 XX Human beta 2 integrin alpha d subunit.
 DE Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW psoriasis; lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 17..1108
 FT Domain

```

FT      /label= Extracellular_domain
FT      150..352
FT      /note= "region homologous to the I (insertion) domain
FT      common to CD11a, CD11b and CD11c"
FT      Domain
FT      1109..1128
FT      /label= Transmembrane_domain
FT      /note= "homologous to the human CD11c transmembrane
FT      region"
FT      1129..1161
FT      /label= Cytoplasmic_domain
FT      Domain
FT      WO9731099-A1.
XX      28-AUG-1997.
XX      24-FEB-1997; 9TWO-US002713.
XX      22-FEB-1996; 9GUS-00605672.
XX      (ICOS-) ICOS CORP.
XX      Gallatin WM, van Der Vieren M;
XX      WPI; 1997-435154/40.
XX      N-PSDB; AAT79220.
XX      Hyalidoma 199M and antibody secreted by it - specific for new rat beta2
XX      integrin subunit, useful to detect subunit in cells and modulate its
XX      activity.
XX      Example 5; Page 116-120; 222pp; English.
XX      This polypeptide comprises a novel human beta 2 integrin subunit,
XX      designated alpha d. Its sequence was deduced from a cDNA clone (see
XX      AAT79220) isolated from a spleen cDNA library. Alpha d is involved in
XX      cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
XX      d polypeptides can be expressed in transformed host cells for use in
XX      assays for identifying antibodies or other compounds that modulate alpha
XX      d activity or which modulate the interaction between alpha d and a
XX      ligand, for treating or preventing diseases in which macrophages are
XX      implicated. Treatment is applicable to disease states in which alpha d
XX      binding, or localised accumulation of cells which express alpha d, is
XX      implicated such as type I diabetes, atherosclerosis, multiple
XX      sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX      distress syndrome and rheumatoid arthritis
XX      Sequence 1161 AA;
XX      Query Match          57.8%; Score 3401; DB 2; Length 1161;
XX      Best Local Similarity 59.3%; Pred. No. 1e-275;
XX      Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
XX      1 FNLDTENAMTFQENARGFGQSVVQLQGRVYVVGAPQEIIVAAHQGLSYQCYSTGSCPEI 60
XX      17 FNLVDEEPTIFQEDAGFGQSVVQFGGSLVVGAPLEVVAANQTCRLYDCAATGMCQPI 76
XX      61 RLQVPVAVNMSIGLSLAATTSPPQLLAGPTVHQTCSNTYVYKGLCELPNSLRQOQPK 120
XX      77 PLHIRPEAVNMSIGLTLAASTNGSLRLLACGPTLHVRQGENSYSGKSCLLLSRW-BIIQT 135
XX      121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTVMEQLKSKTILFSIMQYSEEF 180
XX      136 VPDATPECHQEMDIFVLIDGSGSIDQDNDFNMKGFPVQVMQFEGDTLIPALMQYNLL 195
XX      181 RIHPTKQFQNNPNRSLIKPTOLLGRTHATGKRVKRVRELFTNITGARKQAFKILILL 240
XX      196 KIHFTTQPTSPSQSLVDPIVQLKGLTFTATGILTVVTQLFHKGNGARKSAKKILIVI 255
XX      241 TDEKFGDPLGVEDYVPELDRGCVIRYVIGVDARFSEKSRQELNIVASKPPDRHVFQIN 300
XX      256 TDQCKYKDFLEYSVDVLPQEKAGIIRYALGVGHACGPTARQELNTISSAPPQDHVKVD 315
XX      301 NFPEALXTIQNLRKEIFAIEGTQGTGSSSFHEHMSQEGFSAATISNGPLLSTVGYDWAQ 360

```

```

Db      316 NFPAALASIQKQOEKIYAVDEGTQSRASSFPQHEMSQEGFSTALTMGDLFGAVGSRWSG 375
Qy      361 GVFLYTSKEKSTFINNTRVDSMDNDAYLVAAAIILRNVRQSVLIGAPRYOHIGLVAMFR 420
Db      376 GAFLYPNMSPPTINMSQENVDNRDSYLGSTELALWKGQNLVLAGAPRYOHTGKAVFT 435
Qy      421 QNTGMESNANVKGTQIGAFYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGQGVSCPL 480
Db      436 QVSRQWKKARVYGTQIGSYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGQGVSCPL 495
Qy      481 PRQQRARWCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTVAIGAPEGEDNRGAVYLF 540
Db      496 PRGORVQWQCDVLRGEGQHPWGRFGAALTVLGDVNGDKLTVAIGAPEGEDNRGAVYLF 555
Qy      541 HGTSGSGISPSHSORISAGSLSPRLQYFGOSLGGSGDLTMGDLVDTLVGAQGVHLLRSQ 600
Db      556 HGASESGISPSHSORISAGSLSPRLQYFGOSLGGSGDLTMGDLVDTLVGAQGVHLLRSQ 615
Qy      601 PVLVRVAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREBQIQSWVT 660
Db      616 PVLKVGVMERFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
Qy      661 YDLALDSGRPHSRAVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLBNF 720
Db      674 FDLALDSGRPHSRAVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLBNF 733
Qy      721 SLVGTPLSAFQNLRPVLAEDAQRLFTALPFFKKNCGNDNICQDDLSITF9FMSLDCLVWG 780
Db      734 SLVREPIPSQNLRPVLAQVGSQDLFTASLFFKKNCGQDGLCEGDLGVTLSFGSLQTLTVG 793
Qy      781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSORSWRLACESASTEV 840
Db      794 SSLELVIVTVNAGDSYGTVSVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PIED 852
Qy      841 SGALKSTSCSINPIRPIENSEVTFNITFDYDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db      853 EG-LRSRCSVNPPIFEGNGCFIVTFDYSYKATLDGRLMRASASNNKASSSRATP 911
Qy      901 QLELPVYAYVMVTVSHGVTYKYNLF-TASENTSRVWQYQVSNLQORSILPISLVLPV 959
Db      912 QLELPVYAYVMVTVSHGVTYKYNLF-TASENTSRVWQYQVSNLQORSILPISLVLPV 971
Qy      960 VRLNQTVMDRPOVTFPSNLSSTCHTKERLPSPSHSDFLAELRKAPVNVCSIAVCORICDI 1019
Db      972 VLLNGVAVDVMVMEAPSQL--PCVSEKPPQSHSDFLTQISRSPMLDCSIADCLQPRCDV 1029
Qy      1020 PFFGIOEFNATLKGNLSDFWYIYKTSNHLIIYSTAEILFNDSVFTLLPQOGAFVRSQTE 1079
Db      1030 PSFVSQSELDFTLKGNLSPGVRBTLQKVLVYVSVABITFTDTSVYSQLPGCEAPMRAQME 1089
Qy      1080 TKVEPPFVNPPLIVGSSVGGILLIITLITAAALYKLGFEKQYKQMMSE 1128
Db      1090 MVLDEESVYNAIPIMGSSVGALELLALITATLYKLGFFKRRHYKEMLED 1138
XX      RESULT 15
XX      AAW57491
XX      ID AAW57491 standard; protein; 1161 AA.
XX      AC AAW57491;
XX      DT 24-AUG-1998 (first entry)
XX      DE Human beta2 integrin alpha subunit (alpha d) polypeptide.
XX      KW Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;
XX      reporter-transactivator construct; arteriosclerosis; atherosclerosis;
XX      inflammatory bowel disease; arthritis; multiple sclerosis.
XX      OS Homo sapiens.
XX      Key Location/Qualifiers
XX      FH

```

```

Peptide      1..16
/Note= "signal peptide"
Protein      17..1161
/Note= "mature protein"
Domain      17..1108
/Note= "extracellular domain"
Domain      150..352
/Note= "I (insertion) domain"
Region      1109..1128
/Note= "transmembrane region"
Domain      1129..1161
/Note= "cytoplasmic domain"

US576850-A.
16-JUN-1998.
21-DEC-1994; 94US-00362652.
23-DEC-1993; 93US-00173497.
05-AUG-1994; 94US-00286889.

XX (ICOS-) ICOS CORP.
XX Van Der Vieren M, Gallatin WM;
XX WPI; 1998-361678/31.
XX N-PSDB; RAV31540.
XX Isolation of DNA encoding protein that binds to integrin subunit - using
XX recombinant cells containing reporter-transactivator construct.
XX Example 5; Col 45-54; 86pp; English.
XX This represents a human beta2 integrin alpha subunit (alpha d)
XX polypeptide. This is used in the methods of the invention for isolating a
XX polynucleotide encoding a protein that binds to alpha d. The method
XX comprises transforming or transfecting host cells with a DNA construct
XX comprising a reporter gene under the control of a promoter regulated by a
XX transcription factor having a DNA-binding domain and an activating
XX domain. A first hybrid DNA sequence encoding a fusion of at least part of the
XX alpha d and either the DNA-binding domain or the activating domain of the
XX transcription factor and a library of hybrid DNA sequences encoding
XX fusions of at least part of putative alpha d-binding proteins and the DNA
XX -binding domain or the activating domain of the transcription factor
XX which is not incorporated in the first fusion are expressed in the host
XX cells. The binding of an alpha d-binding protein to alpha d in a
XX particular host cell is detected by determining production of the
XX reporter gene product in the cell. The hybrid DNA sequence encoding the
XX alpha d-binding protein can be isolated from the cell. Alpha d may be
XX useful for treating graft arteriosclerosis, atherosclerosis, diabetes,
XX inflammatory bowel disease, arthritis and multiple sclerosis

SQ Sequence 1161 AA;
Query Match 57.8%; Score 3401; DB 2; Length 1161;
Best Local Similarity 59.3%; Ered. No. 1e-275;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDTEAMTFQENARGFGQSVVQSGSRVVGAPQEIIVANQGRSLYQCDSYSGCEPI 60
DB 17 FNLDVSEPTTFQEDAGGFGQSVVQSGSRVVGAPLEVVANQTCRLYDCAATGMCQPI 76
QY 61 RLOVPEAVNMSLGLSLAATTSPPQLACGPTVHTCSENTVYKGLCFELFGSNLRQPOK 120
DB 77 PLHIRPANNVLSGLHAASTNSRLACGPTLHRCVGENSYSGKSCILLGSRW-EIIOT 135
QY 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFREKMEWVSTVMEQLKSKTLFSLMYSEEP 180
DB 136 VPDATPCRPQEMDIVFLIDGSGSIDQDNFNQKGFQVAVMGQFEGTDTLFALMYSNLL 195
QY 181 RIHPTPEFQNNPRSLIKPITQLGRTHATGRKVVRELENTNGARKNAKILFLL 240

```

```

DB 196 KIHFTFTQFTSPSQSLVDPVQLKGLTFTATGILTVVTLFHHKNGARSAKILLIVI 255
QY 241 TDGEKFGDPLGYEDVIRBELDREGVIRYVIVGSDAFPRSEKSOELMTVASKPRPHVFOIN 300
DB 256 TDGQKYKDPLEYSVDVIPAERKAGIIRYAIYGVGHAFQGPATQELMTISSAPPQDFVVD 315
QY 301 NPEALKTIQNLREKIFAIEGTQTCSSSFHEMSQEGFSAAITSNGPLLSITVGSYDWAG 360
DB 316 NPAALGSIQKLOEKIYAVEGTQSRASSSPQHEMSQEGFSTALTMDGLFLGAVGSFWSG 375
QY 361 GVPLTYSKESKSTFINWTRVDSMDNDAYLGYAAAIIILNRVQSLVILGAPRYOHIGLVAFR 420
DB 376 GAPLYFPNMSPTFINNSENVDMDRSYLGYSYELATWGVQVQLVILGAPRYOHTKAVIT 435
QY 421 QNTGMMESNANVKGTOIGAYFCASLCSVDVDSNGSTDLVLICABHYTEQTRGGQVSVCP 480
DB 436 QVSRQWRKKAETVTOIGSYFCASLCSVDVDSNGSTDLVLICABHYTEQTRGGQVSVCP 495
QY 481 PRGORARWOCDAVLYGEOGPWRGALTVLGVNKGKLTDAVAGPREDNEGAVYLF 540
DB 496 PRGORVWOCDAVLYGEOGPWRGALTVLGVNKGKLTDAVAGPREDNEGAVYLF 555
QY 541 HGTSGSISPSHSQRIRAGSKLSPRLQYFQSLSGQDLTMDGLVDLTVGAQHVLRLRSQ 600
DB 556 HGASESGISPSHSQRIRAGSKLSPRLQYFQSLSGQDLTMDGLVDLTVGAQHVLRLRSQ 615
QY 601 PVLKVAIMEPNPREVARNVFECDQVVKGEAGEVRLVOKSTRDRLEGOIQSVVT 660
DB 616 PVLKVGVAWRFVVEKAVYRCWEEKPSALRAGDATVCLTIQKSSLDOL--GDIQSVR 673
QY 661 YDLALDSGPHSRAVFNETKSTRQTVLGTOTCEPLKLQLPNCIEDPVSPIVLRNF 720
DB 674 FDLALDPGLTSRAIFNETKSTRQTVLGTOTCEPLKLQLPNCIEDPVSPIVLRNF 733
QY 721 SLVGPPLSAFGLNRPVLAEDAORLFTALPPEKKNAGNDNICODDLSITFSFMSLDCLVVG 780
DB 734 SLVREPLSPQNLRLVAVGSDLTASLPPEKKNAGNDNICODDLSITFSFMSLDCLVVG 793
QY 781 GPREFNVTYVNDGEDSYRQVTFPPPLDLSYKRVSTLQNRQSRWRACBSASSTEV 840
DB 794 SLELNVIYVWVNDGEDSYRQVTFPPPLDLSYKRVSTLQNRQSRWRACBSASSTEV 852
QY 841 SGALKSTCSINHPFPENSEVTFNITPDVDSKASLGNKLLKANVTSENNMPTNKTET 900
DB 853 EG-LRSSRCSVNHPIFHBGSGNGTPIVTFDVSYKATLGDRLMRASSENKASSSKATP 911
QY 901 QLELPKYAVVWVTSHGVSYTKYLP-TASENTRVWQHOYQVSNLQORSILPILVLP 959
DB 912 QLELPKYAVVWVTSHGVSYTKYLP-TASENTRVWQHOYQVSNLQORSILPILVLP 971
QY 960 VRINQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVAVNCSIAVCORICDI 1019
DB 972 VLLNGVAVWVWMEAPSQSL--PCVSEKPKPHSDFLTQISRSPLMDCSIADCLQPRCDV 1029
QY 1020 PFGIOEENATLKNLSFDWYIKTSHNHLIVSTAELFENDSVFTLLPGOGAFVRSOTE 1079
DB 1030 PFSVQOEELDFTLKGLNLSFGWRETLOKQVILVSVABITFDTSVYSQLPQOEAPRAQME 1089
QY 1080 TKVEPFEVNPULPLVSGSVGGLLLALITAAALYKLGFFKRYQKDMSE 1128
DB 1090 MYLEDEEVNAPIIMGSSVAGALLALITATLYKLGFFKRYKHEMLED 1138

```

Search completed: June 7, 2004, 17:11:39
Job time : 53.4608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds

(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKRQYKDMWSEGGPGABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5884	100.0	1137	10	US-09-902-481A-4
2	5857	99.5	1137	10	US-09-902-481A-3
3	5852	99.5	1153	9	US-09-350-259-3
4	5852	99.5	1153	10	US-09-902-481A-1
5	5852	99.5	1153	10	US-09-891-943-3
6	5852	99.5	1153	14	US-10-144-259-30
7	5852	99.5	1153	14	US-10-207-655-176
8	5851	99.4	1137	10	US-09-902-481A-5
9	5845	99.3	1137	10	US-09-902-481A-6
10	5836.5	99.2	1152	9	US-09-945-265-4
11	3459	58.8	1163	15	US-10-116-275-204
12	3436	58.4	1163	9	US-09-350-259-4
13	3436	58.4	1163	10	US-09-891-943-2
14	3401	57.8	1161	9	US-09-350-259-2
15	3401	57.8	1161	10	US-09-891-943-2

16 3385.5 57.5 1161 9 US-09-350-259-99 Sequence 99, Appl
17 3385.5 57.5 1161 10 US-09-891-943-99 Sequence 99, Appl
18 3224.5 54.8 1161 9 US-09-350-259-55 Sequence 55, Appl
19 3224.5 54.8 1161 10 US-09-891-943-55 Sequence 55, Appl
20 3212 54.6 1161 9 US-09-350-259-53 Sequence 53, Appl
21 3212 54.6 1161 10 US-09-891-943-53 Sequence 53, Appl
22 3203.5 54.4 1151 9 US-09-350-259-37 Sequence 37, Appl
23 3203.5 54.4 1151 10 US-09-891-943-37 Sequence 37, Appl
24 3191 54.2 1155 9 US-09-350-259-46 Sequence 46, Appl
25 3191 54.2 1155 10 US-09-891-943-46 Sequence 46, Appl
26 1848 31.4 369 12 US-10-087-192-1212 Sequence 1212, Appl
27 1547.5 26.3 1170 9 US-09-945-265-2 Sequence 2, Appl
28 1543.5 26.2 1170 12 US-10-261-164-1 Sequence 1, Appl
29 1520 25.8 1223 16 US-10-408-765A-295 Sequence 295, Appl
30 1350.5 23.0 1086 16 US-10-408-765A-1871 Sequence 1871, Appl
31 1229.5 20.9 494 9 US-09-350-259-103 Sequence 103, Appl
32 1229.5 20.9 494 10 US-09-891-943-103 Sequence 103, Appl
33 1141.5 19.4 413 9 US-09-350-259-101 Sequence 101, Appl
34 1141.5 19.4 413 10 US-09-891-943-101 Sequence 101, Appl
35 1140 19.4 1179 14 US-10-177-550-2 Sequence 2, Appl
36 1140 19.4 1179 15 US-10-173-551-2 Sequence 2, Appl
37 1093.5 18.6 1151 10 US-09-894-130-103 Sequence 103, Appl
38 1093.5 18.6 1151 10 US-09-836-353A-103 Sequence 103, Appl
39 1093.5 18.6 1179 12 US-09-318-715-250 Sequence 250, Appl
40 1092 18.6 1188 15 US-10-291-265-810 Sequence 810, Appl
41 1088 18.5 1188 15 US-10-291-265-338 Sequence 338, Appl
42 1084.5 18.4 1189 10 US-09-984-130-35 Sequence 35, Appl
43 1084.5 18.4 1189 10 US-09-836-353A-35 Sequence 35, Appl
44 1084.5 18.4 1189 12 US-10-262-839-4 Sequence 4, Appl
45 1074 18.3 589 12 US-10-261-164-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-902-481A-4

; Sequence 4, Application US/09902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimaoka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/902.481A

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,600

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-4

Query Match 100.0%; Score 5884; DB 10; Length 1137;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGSVOLQSGRVVVGAPQIIVANQSGSLVQCDYSTGSCPEI 60

Db 1 FNLDTENAMTQENARGFGSVOLQSGRVVVGAPQIIVANQSGSLVQCDYSTGSCPEI 60

Qy 61 RLQVPVAVNMVSLGSLAAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFQSNLRQQPQK 120

Db 61 RLQVPVAVNMVSLGSLAAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFQSNLRQQPQK 120

Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKWSTVWEOQLKSKTLLSLWQYSSEF 180


```
121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTLFSLMOYSEEP 180
181 RIHFTKPEFONNPNPSLSIKPITQLGRTHATGLKRVVRELFNITNGARKNAKPLIFLL 240
181 RIHFTKPEFONNPNPSLSIKPITQLGRTHATGLKRVVRELFNITNGARKNAKPLIFLL 240
241 TDEKFGDPLGYEDVPELDREGVIRYVGVGDAFSEKSRBELNTVASKPRDHVFOIN 300
241 TDEKFGDPLGYEDVPELDREGVIRYVGVGDAFSEKSRBELNTVASKPRDHVFOIN 300
301 NPEALKTIONQUREKIPIAIEGTGTGSSSSEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
301 NPEALKTIONQUREKIPIAIEGTGTGSSSSEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR 420
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVCP 480
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVCP 480
481 PRGORAEWOCDAVLYGEOGQOPWGRFGAALTVDVNGDXLTDVAIGARGEEDNRGAVILP 540
481 PRGORAEWOCDAVLYGEOGQOPWGRFGAALTVDVNGDXLTDVAIGARGEEDNRGAVILP 540
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPEKNCNDNICQDDLSITPFSMSLCLVVG 780
721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPEKNCNDNICQDDLSITPFSMSLCLVVG 780
781 GRPEFNVTVVRNDGSDVYRTQVTPFPPLDLSVRKYSTLQNRORSORSLACASSTEV 840
781 GRPEFNVTVVRNDGSDVYRTQVTPFPPLDLSVRKYSTLQNRORSORSLACASSTEV 840
841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLFLVPV 960
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLFLVPV 960
961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISLAVCOIRIQC 1020
961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISLAVCOIRIQC 1020
1021 PFGIQEENFATLKNLSFPWYIKTSHNLLIYVSTABILFNDSPVFLPQOGAFVRSOTET 1080
1021 PFGIQEENFATLKNLSFPWYIKTSHNLLIYVSTABILFNDSPVFLPQOGAFVRSOTET 1080
1081 KVEPPEFVNPPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKDMWMSGGPPGABFPQ 1137
1081 KVEPPEFVNPPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKDMWMSGGPPGABFPQ 1137
```

RESULT 2

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.5%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY	1	FNLDTENAMTPOENARGFGQSVQLOQSRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI	60
DB	1	FNLDTENAMTPOENARGFGQSVQLOQSRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI	60
QY	61	RLQVPVEAVNNSLGLSLAATTSPQILLACGPTVHTQTCENTYVKGCLCPLFGSNLRQQPK	120
DB	61	RLQVPVEAVNNSLGLSLAATTSPQILLACGPTVHTQTCENTYVKGCLCPLFGSNLRQQPK	120
QY	121	FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTLFSLMOYSEEP	180
DB	121	FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTLFSLMOYSEEP	180
QY	181	RIHFTKPEFONNPNPSLSIKPITQLGRTHATGLKRVVRELFNITNGARKNAKPLIFLL	240
DB	181	RIHFTKPEFONNPNPSLSIKPITQLGRTHATGLKRVVRELFNITNGARKNAKPLIFLL	240
QY	241	TGSKFGDPLGYEDVPELDREGVIRYVGVGDAFSEKSRBELNTVASKPRDHVFOIN	300
DB	241	TGSKFGDPLGYEDVPELDREGVIRYVGVGDAFSEKSRBELNTVASKPRDHVFOIN	300
QY	301	NFEALKTIONQUREKIPIAIEGTGTGSSSSEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG	360
DB	301	NFEALKTIONQUREKIPIAIEGTGTGSSSSEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG	360
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR	420
DB	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAILLRNVQSLVLCAPRYQHIGLVAMFR	420
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVCP	480
DB	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVCP	480
QY	481	PRGORAEWOCDAVLYGEOGQOPWGRFGAALTVDVNGDXLTDVAIGARGEEDNRGAVILP	540
DB	481	PRGORAEWOCDAVLYGEOGQOPWGRFGAALTVDVNGDXLTDVAIGARGEEDNRGAVILP	540
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ	600
DB	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ	600
QY	601	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
DB	601	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
QY	661	YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	720
DB	661	YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	720
QY	721	SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPEKNCNDNICQDDLSITPFSMSLCLVVG	780

Db 721 SLVGTPLSAFGLNRLPVLAEDAQRLETFALFPFKKNCNDNIQODDLSTIFSFMSLDCLVVG 780
Qy 781 GPRFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Db 781 GPRFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Qy 901 QLELPVKAVYVMVTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLFLVPV 960
Db 901 QLELPVKAVYVMVTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLFLVPV 960
Qy 961 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Db 961 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Qy 1021 FFGIOBEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Db 1021 FFGIOBEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Qy 1081 KVEPPEVFNPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137
Db 1081 KVEPPEVFNPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1ei Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.5%; Score 5852; DB 9; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENATFOENARGFGQSVVQLQSGRVVVGAPQEIIVAANQKSLYOCYSTGSCPEI 60
Db 17 FNLDTENATFOENARGFGQSVVQLQSGRVVVGAPQEIIVAANQKSLYOCYSTGSCPEI 76
Qy 61 RLQVPVAVNNSLGLSLAATSPOLLACGPTVHTQCSNTYVVKGLCFPGSNLRQOPK 120
Db 77 RLQVPVAVNNSLGLSLAATSPOLLACGPTVHTQCSNTYVVKGLCFPGSNLRQOPK 136
Qy 121 FPEALRGCPQSDSDAFILIDSGSIIIPHDPRMKKSWTVWBLKSKTSLPSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDAFILIDSGSIIIPHDPRMKKSWTVWBLKSKTSLPSLMQYSEEF 196

Qy 181 RIHPTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFLL 240
Db 197 RIHPTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFLL 256
Qy 241 TDEKPKDPLGYEDVIBELDRREGVIRVYIGVDAFRSEKSRQELNVTASPPRDRHVQIN 300
Db 257 TDEKPKDPLGYEDVIBELDRREGVIRVYIGVDAFRSEKSRQELNVTASPPRDRHVQIN 316
Qy 301 NFEALTKIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATNSNGPLLSITVGSYDWA 360
Db 317 NFEALTKIQNLREKIPAIETGTGSSSSPEHEMSQEGFSAATNSNGPLLSITVGSYDWA 376
Qy 361 GVFLYTSKESKSTFNNTRVDSNDNDVILGYAAAIILRNRVQSLVGLGAPRHOHIGLVAMER 420
Db 377 GVFLYTSKESKSTFNNTRVDSNDNDVILGYAAAIILRNRVQSLVGLGAPRHOHIGLVAMER 436
Qy 421 QNTGMESSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYTYEOTRGQVSVCP 480
Db 437 QNTGMESSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYTYEOTRGQVSVCP 496
Qy 481 PRQARWQCDVILYGEQGFQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVL 540
Db 497 PRQARWQCDVILYGEQGFQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVL 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMBENPREVARNVFEQNDQVVKKEAGEVAVCLHVQKSTDRLEREGIQSVVT 660
Db 617 PVLRVKAIMBENPREVARNVFEQNDQVVKKEAGEVAVCLHVQKSTDRLEREGIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRLPVLAEDAQRLETFALFPFKKNCNDNIQODDLSTIFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRLPVLAEDAQRLETFALFPFKKNCNDNIQODDLSTIFSFMSLDCLVVG 796
Qy 781 GPRFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 840
Db 797 GPRFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 916
Qy 901 QLELPVKAVYVMVTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLFLVPV 960
Db 917 QLELPVKAVYVMVTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSISLFLVPV 976
Qy 961 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036
Qy 1021 FFGIOBEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSPVTLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSPVTLPGQAFVRSQTET 1096
Qy 1081 KVEPPEVFNPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137
Db 1097 KVEPPEVFNPLPIVGVSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1153

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia

APPLICANT: Mayo, Stephen
TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
FILE REFERENCE: A-70586-1/FT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: mat peptide
LOCATION: (17)..()
OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLVQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLVQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQQPQK 136
QY 121 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEF 180
DB 137 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TGEKFGDPLGVEDVPIPELDREGVIRYVIGVDAPRSEKSRQRLNTVASKPPRDRHVQIN 300
DB 257 TGEKFGDPLGVEDVPIPELDREGVIRYVIGVDAPRSEKSRQRLNTVASKPPRDRHVQIN 316
QY 301 NFEALATIONLREKIFALEGTQSGSSFEHMSQEGFSAITNSGPLLSVGVSDWAG 360
DB 317 NFEALATIONLREKIFALEGTQSGSSFEHMSQEGFSAITNSGPLLSVGVSDWAG 376
QY 361 GVFLYTSREKSTFINNTRVDSMDNDAYLGAYAAAILLRNVQSLVLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSREKSTFINNTRVDSMDNDAYLGAYAAAILLRNVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGSLCSVDVDSNGSTDLLVLGAPHYVEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGSLCSVDVDSNGSTDLLVLGAPHYVEOTRGQVSVCP 496
QY 481 PRGORARWOCDAVLYGEGQOPWCRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORARWOCDAVLYGEGQOPWCRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQRAGSKLSPRLQYFGOSLGGQDLWDGLVDLTVCAGQHVILLRSQ 600
DB 557 HGTSGSGISPSHQRAGSKLSPRLQYFGOSLGGQDLWDGLVDLTVCAGQHVILLRSQ 616
QY 601 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSAVFNETHKSTRTOTVGLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVFNETHKSTRTOTVGLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736
QY 721 SLVGTPLSARGNLRPVLAEDAQRLLFTALFPFEKNCNGNDNICQDDLSTITPSFMSLDCLVVG 780

DB 737 SLVGTPLSARGNLRPVLAEDAQRLLFTALFPFEKNCNGNDNICQDDLSTITPSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYEKVTSLQNSORSRWRLACBSASSTEV 840
DB 797 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYKRVSTLQNSQRSRWRLACBSASSTEV 856
QY 841 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTPEF 900
DB 857 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTPEF 916
QY 901 QLELPVKYAVYVMVTSKGVSTKYLNFTASENTSRVMQHOYVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSKGVSTKYLNFTASENTSRVMQHOYVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNCIAVCQRIQCDIP 1036
QY 1021 PFGIOBEFNATLKNLSFDWYIKTSHHLLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1037 PFGIOBEFNATLKNLSFDWYIKTSHHLLIVSTABILFNDVSFTLLPGOGAFVRSOTET 1096
QY 1081 KVPEPEVNPPLPLIVGSSVGGLLHLLALITAAALYKLGFFKQYKDMNSEGGPPGAEPO 1137
DB 1097 KVPEPEVNPPLPLIVGSSVGGLLHLLALITAAALYKLGFFKQYKDMNSEGGPPGAEPO 1153
RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLVQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLVQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQQPQK 136
QY 121 PPEALRGCPQSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEF 180
DB 137 PPEALRGCPQSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEF 196
QY 181 RHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256

```
241 TDGEKFGDPLGYEDVPELDREGVIRVIGVDGAFRSEKSRQELNVTASKPPRDRHVPQIN 300
Db TDGEKFGDPLGYEDVPELDREGVIRVIGVDGAFRSEKSRQELNVTASKPPRDRHVPQIN 316
301 NFEALKTIONLRKIFALEGTOTGSSSFHEMSQGFSAITNSGPLLSITVGSYDWAG 360
Db NFEALKTIONLRKIFALEGTOTGSSSFHEMSQGFSAITNSGPLLSITVGSYDWAG 376
361 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 420
Db GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 436
421 QNTGWBESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 480
Db QNTGWBESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 496
481 PRGORARWQCDVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db PRGORARWQCDVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
601 PVLRVKAIMBNPREVARNVFECDNDQVVKKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db PVLRVKAIMBNPREVARNVFECDNDQVVKKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCEIDPSPVILRLNF 720
Db YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCEIDPSPVILRLNF 736
721 SLVGTPLSAFCNLRPVLAEDAORLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db SLVGTPLSAFCNLRPVLAEDAORLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFEFPDLDSYRKVSTLQNRQSRWELACESASSTEV 840
Db GPREFNVTVVRNDGEDSYRTQVTFEFPDLDSYRKVSTLQNRQSRWELACESASSTEV 856
841 SGALSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db SGALSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
901 QLELPVKAVYVMTSHGVSTKYLNTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960
Db QLELPVKAVYVMTSHGVSTKYLNTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 976
961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAHLRKAQVNCIAVCORIQCDDIP 1020
Db RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAHLRKAQVNCIAVCORIQCDDIP 1036
1021 FPGIOEFNATLKGNLSPDWYIKTSHNLLIVSTABILFNDVFTLLPQCGAFVRSQTET 1080
Db FPGIOEFNATLKGNLSPDWYIKTSHNLLIVSTABILFNDVFTLLPQCGAFVRSQTET 1096
1081 KVEPFEPVNPFLPIVGVSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPGABQP 1137
Db KVEPFEPVNPFLPIVGVSSVGGILLALITAAALYKLGFFKQYKQKDMWSEGGPGABQP 1153
```

RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
```

```
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 FNLDENAMTFQENARFGQSVVQLQSSRVVVGAPQEIIVAAQNGSSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTFQENARFGQSVVQLQSSRVVVGAPQEIIVAAQNGSSLYQCDYSTGSCPEI 76
61 RLOVPVEAVNMSLGLSLAATTPPQLLACQPTVHQTCSNTYVVKGLCFPLFGSNLRQOPQX 120
Db 77 RLOVPVEAVNMSLGLSLAATTPPQLLACQPTVHQTCSNTYVVKGLCFPLFGSNLRQOPQX 136
121 FPBALRGCPQEDSDIAFLIDGSGSIIPHPDRRKEWVSTWEOQKKSKTLFSLMOYSEEF 180
Db 137 FPBALRGCPQEDSDIAFLIDGSGSIIPHPDRRKEWVSTWEOQKKSKTLFSLMOYSEEF 196
181 RIHFTFKFQNNPNPSLIPITQLGRHTATGLRAVVELEFNI TNGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNPSLIPITQLGRHTATGLRAVVELEFNI TNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVPELDREGVIRVIGVDGAFRSEKSRQELNVTASKPPRDRHVPQIN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRVIGVDGAFRSEKSRQELNVTASKPPRDRHVPQIN 316
301 NFEALKTIONLRKIFALEGTOTGSSSFHEMSQGFSAITNSGPLLSITVGSYDWAG 360
Db 317 NFEALKTIONLRKIFALEGTOTGSSSFHEMSQGFSAITNSGPLLSITVGSYDWAG 376
361 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 436
421 QNTGWBESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 480
Db 437 QNTGWBESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 496
481 PRGORARWQCDVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
601 PVLRVKAIMBNPREVARNVFECDNDQVVKKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIMBNPREVARNVFECDNDQVVKKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCEIDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQTCETLKLQPNCEIDPSPVILRLNF 736
721 SLVGTPLSAFCNLRPVLAEDAORLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFCNLRPVLAEDAORLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFEFPDLDSYRKVSTLQNRQSRWELACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFEFPDLDSYRKVSTLQNRQSRWELACESASSTEV 856
```


Query Match	99.4%;	Score 5851;	DB 10;	Length 1137;
Seet Local Similarity	99.3%;	Pred. No. 0;		
Matches 1129;	Conservative	5;	Mismatches	3;
			Indels	0;
			Gaps	0;
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSVVVVGAPQEIIVAAQNRGSLYQDYSTGSCBPI	60	
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSVVVVGAPQEIIVAAQNRGSLYQDYSTGSCBPI	60	
Qy	61	RLOVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFLFGSNLROPOK	120	
Db	61	RLOVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFLFGSNLROPOK	120	
Qy	121	FPFALGCGQEDSDIAFLIDGSGIIPHDFRRKMWSTVMEOLKSKKTLFSLMOYSEEF	180	
Db	121	FPFALGCGQEDSDIAFLIDGSGIIPHDFRRKMWSTVMEOLKSKKTLFSLMOYSEEF	180	
Qy	181	RIHFTPKFQNNPNRSLIKPIITOLLGRTHATGLRKVVRLEPNIITNGARKQNAFKILFL	240	
Db	181	RIHFTPKFQNNPNRSLIKPIITOLLGRTHATGLRKVVRLEPNIITNGARKQNAFKILIL	240	
Qy	241	TGQEKFGDPLGYEDVISELDREGVIRVVGVPAPRSEKSRQELNIVASKPPRHHVQQLN	300	
Db	241	TGQEKFGDPLGYEDVISELDREGVIRVVGVPAPRSEKSRQELNIVASKPPRHHVQQLN	300	
Qy	301	NFEALXTIQOLREKIFAIBGTGTGSSSSPEHEMSQBFSAATISNGPLSTVGSYDWAG	360	
Db	301	NFEALXTIQOLREKIFAIBGTGTGSSSSPEHEMSQBFSAATISNGPLSTVGSYDWAG	360	
Qy	361	GVFLYTSKEKSTINTRVDSNDADYLGAAAIILNRVQSLVLCAPRVOHGLGVAMPR	420	
Db	361	GVFLYTSKEKSTINTRVDSNDADYLGAAAIILNRVQSLVLCAPRVOHGLGVAMPR	420	
Qy	421	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSITDLVLI GAPHYEQTGCGQSVCP	480	
Db	421	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSITDLVLI GAPHYEQTGCGQSVCP	480	
Qy	481	PGORARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAI GAPGEDNRGAVILP	540	
Db	481	PGORARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAI GAPGEDNRGAVILP	540	
Qy	541	HGTSGSGISPSHQRISAGSKLSPLOYFGOSLGGQDLTMDGLVDLITVGAQGHVLLRSQ	600	
Db	541	HGTSGSGISPSHQRISAGSKLSPLOYFGOSLGGQDLTMDGLVDLITVGAQGHVLLRSQ	600	
Qy	601	PVLRVKAIIMEFNPREVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDLREGQISVVT	660	
Db	601	PVLRVKAIIMEFNPREVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDLREGQISVVT	660	
Qy	661	YDLALDSGRPHSRVAFNETKNSIRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	720	
Db	661	YDLALDSGRPHSRVAFNETKNSIRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	720	
Qy	721	SLVGTPLSAFGLNRPVLAEADAQRLLFTALPFPEKNCNDNICQDDLSITTFSPMSLDCI	780	
Db	721	SLVGTPLSAFGLNRPVLAEADAQRLLFTALPFPEKNCNDNICQDDLSITTFSPMSLDCI	780	
Qy	781	GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	840	
Db	781	GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	840	
Qy	841	SGALKSTSCSINHPIFPENSEVTENIFPDVDSKASLGNKLLKANVTSENNMPRTNKT	900	
Db	841	SGALKSTSCSINHPIFPENSEVTENIFPDVDSKASLGNKLLKANVTSENNMPRTNKT	900	
Qy	901	QLELPVKYAVTVMTVSHGVSTKYILNFTAGENTSIRVMOHQVSNLQORSIPISLVLVPPV	960	
Db	901	QLELPVKYAVTVMTVSHGVSTKYILNFTAGENTSIRVMOHQVSNLQORSIPISLVLVPPV	960	
Qy	961	RINQTVIWRDPQVTFSENLISSTCTCKEKLPSHSDFLAEELKAPVNVCSIAVCQRIQCDIP	1020	
Db	961	RINQTVIWRDPQVTFSENLISSTCTCKEKLPSHSDFLAEELKAPVNVCSIAVCQRIQCDIP	1020	
Qy	1021	PFQIEBENFATLKGMLSPDWITKISHNHLIIVSTAFLIFNDVSFTLLPGQAFVRSQTET	1080	

Db	1021	PFGLQEFNATLKENLSFDWYKNTSENHLLIVSTAEILLVNDVSFTLLPQCGAFVRSQYET 1080
Qy	1081	KVBFPEVNPPLPIVGSVSGLLLLALITLAALYKLGFFKQYKDMHSEGGPPGAEQ 1137
Db	1081	KVBFPEVNPPLPIVGSVSGLLLLALITLAALYKLGFFKQYKDMHSEGGPPGAEQ 1137
RESULT 9		
US-09-902-481A-6		
; Sequence 6, Application US/09902481A		
; Publication No. US20030054440A1		
; GENERAL INFORMATION:		
; APPLICANT: Springer, Timothy		
; APPLICANT: Shimooka, Motomu		
; APPLICANT: Shifman, Julia		
; APPLICANT: Mayo, Stephen		
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY		
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK		
; CURRENT APPLICATION NUMBER: US/09/902,481A		
; CURRENT FILING DATE: 2001-07-09		
; PRIOR APPLICATION NUMBER: US 60/216,600		
; PRIOR FILING DATE: 2000-07-07		
; NUMBER OF SEQ ID NOS: 7		
; SOFTWARE: Patent in version 3.1		
; SEQ ID NO 6		
; LENGTH: 1137		
; TYPE: PRT		
; ORGANISM: Artificial sequence		
; FEATURE:		
; OTHER INFORMATION: synthetic		
US-09-902-481A-6		
Query Match 99.3%; Score 5845; DB 10; Length 1137;		
Best Local Similarity 99.0%; Pred. No. 0;		
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	PNLDTENAMTFQENARGFGSGVVOQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db	1	PNLDTENAMTFQENARGFGSGVVOQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Qy	61	RQVPEVAVNMISLGLSLAAITSPDPOLLACGPTVHQTCSNTYKGLCFPLGSLNRQOPQK 120
Db	61	RQVPEVAVNMISLGLSLAAITSPDPOLLACGPTVHQTCSNTYKGLCFPLGSLNRQOPQK 120
Qy	121	FPFALRGCPQSDIAFLIDGSGSIIIPHDPRMKWVSTVWEOLKKKTLFSLMQYSEEF 180
Db	121	FPFALRGCPQSDIAFLIDGSGSIIIPHDPRMKWVSTVWEOLKKKTLFSLMQYSEEF 180
Qy	181	RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRPELFNITNGARKNAKPIFL 240
Db	181	RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRPELFNITNGARKNAKPIFL 240
Qy	241	TGSEKFGDPLGYEYVIFELDEGVIRYVIGVDAPRSEKSQLNTVASKPPRDHVFQIN 300
Db	241	TGSEKFGDPLGYEYVIFELDEGVIRYVIGVDAPRSEKSQLNTVASKPPRDHVFQIN 300
Qy	301	NFEALKTIQNLREKIPALEGTQTGSSSSFEHNSQEGFSAAITNSGPLLSTVGSYDWA 360
Db	301	NFEALKTIQNLREKIPALEGTQTGSSSSFEHNSQEGFSAAITNSGPLLSTVGSYDWA 360
Qy	361	GVFLYTSKEKSTFTNMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db	361	GVFLYTSKEKSTFTNMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Qy	421	QNTGMWESNANVKGTQIGAYFGASILCSVDVDSNGSTDLVLIGAPHYETQTRGGVSV 480
Db	421	QNTGMWESNANVKGTQIGAYFGASILCSVDVDSNGSTDLVLIGAPHYETQTRGGVSV 480
Qy	481	PRGQARVQCDVAVLYEGQGGPWGRFGAALTVLGVDNGDKLTDVAIGAGEEDNREGAVYL 540
Db	481	PRGQARVQCDVAVLYEGQGGPWGRFGAALTVLGVDNGDKLTDVAIGAGEEDNREGAVYL 540


```

541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 600
541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 600
601 PVLAVKAIMBNPREVARNVFECDQVVKGEKAGEVRVCLHVOKSTRDRLEQIQSVVT 660
601 PVLAVKAIMBNPREVARNVFECDQVVKGEKAGEVRVCLHVOKSTRDRLEQIQSVVT 660
661 YDLALDSGRPHSAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
661 YDLALDSGRPHSAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPKNGCNDNICQDDLSITPFSMGLDCLVVG 780
721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPKNGCNDNICQDDLSITPFSMGLDCLVVG 780
781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACESASTEV 840
781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACESASTEV 840
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
901 QLELPVKAYVMVTVSHGVSTKYLNTASENTSRVMQHOXOVSNLQORSLSPIVLPLVP 960
901 QLELPVKAYVMVTVSHGVSTKYLNTASENTSRVMQHOXOVSNLQORSLSPIVLPLVP 960
961 RLMTQVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
961 RLMTQVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
1021 PFGIOEEFNATLKGSLSPDQVYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGAFVRSQTE 1080
1021 PFGIOEEFNATLKGSLSPDQVYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGAFVRSQTE 1080
1081 KVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKDMSEGGPPGABPQ 1137
1081 KVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKDMSEGGPPGABPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-945-265-4

Query Match 99.2%; Score 5936.5; DB 9; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOBIVAANQSGSLYQCDYSTGSCPT 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOBIVAANQSGSLYQCDYSTGSCPT 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 120

```

```

77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 136
121 PPEARLGGCPQEDSDIAFLIDGSGSIIPHDPRMKEMWVSTMEOQLKSKTLPMLQVSEEF 180
137 PPEARLGGCPQEDSDIAFLIDGSGSIIPHDPRMKEMWVSTMEOQLKSKTLPMLQVSEEF 196
181 RIHFTFKSPQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNFITNGARKNAFKILFUL 240
197 RIHFTFKSPQNNPNRSLIKPITQLLGRTHATGLRKVVRRELNFITNGARKNAFKILVVI 256
241 TDGBKFGPLGYEDVTPELDREGVRYVIGVGDAPRSBKSROELNATVASKPRDHVQJLN 300
257 TDGBKFGPLGYEDVTPELDREGVRYVIGVGDAPRSBKSROELNATVASKPRDHVQJLN 316
301 NFEALKTIQNLQREKI PAIEGTQTQSSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWA 360
317 NFEALKTIQNLQREKI PAIEGTQTQSSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWA 376
361 GVFLYTSKEKSTFINKTRVDSMDNDAIYGAAAAIILNRVQSLVAGARYOHI GLVAMPR 420
377 GVFLYTSKEKSTFINKTRVDSMDNDAIYGAAAAIILNRVQSLVAGARYOHI GLVAMPR 436
421 QNTGHWBSNANVKGITQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYBQTRGGQSVVCP 480
437 QNTGHWBSNANVKGITQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYBQTRGGQSVVCP 496
481 PRGQBARQCDAVLVGBGQGPWGRFGAALTVDVYNGDKLTDVAI GAPGEEDNRGAVYLP 540
497 PRG-BARQCDAVLVGBGQGPWGRFGAALTVDVYNGDKLTDVAI GAPGEEDNRGAVYLP 555
541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 600
556 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLSGGDLTMDGLVDLTGVAQGHVLLRSQ 615
601 PVLAVKAIMBNPREVARNVFECDQVVKGEKAGEVRVCLHVOKSTRDRLEQIQSVVT 660
616 PVLAVKAIMBNPREVARNVFECDQVVKGEKAGEVRVCLHVOKSTRDRLEQIQSVVT 675
661 YDLALDSGRPHSAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
676 YDLALDSGRPHSAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 735
721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPKNGCNDNICQDDLSITPFSMGLDCLVVG 780
736 SLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPKNGCNDNICQDDLSITPFSMGLDCLVVG 795
781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACESASTEV 840
796 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACESASTEV 855
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 915
901 QLELPVKAYVMVTVSHGVSTKYLNTASENTSRVMQHOXOVSNLQORSLSPIVLPLVP 960
916 QLELPVKAYVMVTVSHGVSTKYLNTASENTSRVMQHOXOVSNLQORSLSPIVLPLVP 975
961 RLMTQVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
976 RLMTQVIMDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1035
1021 PFGIOEEFNATLKGSLSPDQVYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGAFVRSQTE 1080
1036 PFGIOEEFNATLKGSLSPDQVYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGAFVRSQTE 1095
1081 KVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKDMSEGGPPGABPQ 1137
1096 KVEPPEVNPPLIIVGSSVGGLLLALITAALYKLGFFKQYKDMSEGGPPGABPQ 1152

```

RESULT 11

US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Eran Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 58.8%; Score 3459; DB 15; Length 1163;
Best Local Similarity 60.9%; Pred. No. 4.5e-310;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGSGRVRVVGAPQEIIVAAHQSGSLYQCYDSTGSCBPI 60
DB 20 FNLDTENAMTFOENARGFGQSVVQLGSGRVRVVGAPQEIIVAAHQSGSLYQCYDSTGSCBPI 79

QY 61 RLOVPVEAVNMVSLGSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 120
DB 80 GLQVPEAVNMVSLGSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 137

QY 121 FPEALRGCPQEDSDIAFLDGSISIIIPHDFRMEKWSVTVMEQLKSKTLFSLMOYSEBF 180
DB 138 LPVSRQECRQEQDIIVFLDGSISISRNPFATWVNFRAVISQFQSTQFSLMQFSNKF 197

QY 181 RHFTPEFQNNPNRSLKIPITQLGRTHATGLKRVKURELNIYNGAKNAFKILFLL 240
DB 198 QTHFTPEFRTSNPLSLASVHQLGQFTYATAIQNVVHRLFHASGARRDATKILVI 257

QY 241 TDEKFGDPLGYEDVTPELDREGVIRVIGVGDAPFRSEKSRQELNTVWAKPPRDHVFQIN 300
DB 258 TDCKEGSDLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKELNDIAKSPQSHIFKVE 317

QY 301 NPEALKTIONOLREKIPAIETGTGSSSPHEHMSOEGFSAATSNGLLSTVGSVDWAG 360
DB 318 DFDALKDIOQLKKEKIPAIETGTGSSSPHEHMSOEGFSAATSNGLLSTVGSVDWAG 377

QY 361 GVFLYTSKESKSTINMTRVDSNDNDVGLVYAAAIILNRNRVQSLVGLAPRVQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINNSQENNVDRDYSYLGYSTELALWKGVQSLVGLAPRVQHTGKAVIT 437

QY 421 QNTGMESNANVKTGTGIFGASLCSVDVDSNGSDTLVLIGAPHYVEQTRGQSVQCP 480
DB 438 QVSRQEMKAEVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYVEQTRGQSVQCP 497

QY 481 PRGRARWQCDVLYGEGQGPWRFGAALTIVLGVNCKDKLTVIAGPGEENRGAVYLF 540
DB 498 PRGR-RWQCDVLYGEGQGPWRFGAALTIVLGVNCKDKLTVIAGPGEENRGAVYLF 556

QY 541 HGTSGSGISPSHRSORLAGSKLSPRYFGQSLGCGQDLTMDGLVLTGVAQGHVLLRSQ 600
DB 557 HGVLPISPSHRSORLAGSKLSPRYFGQSLGCGQDLTMDGLVLTGVAQGHVLLRSQ 616

QY 601 PVLVRKAIMFNPREAVNPFECNDQVVKGEAGEVVRVCHVOKSTRDLRBEQIOQSVT 660
DB 617 PVLVWGVSMQFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSVT 676

QY 661 YDLALDSGRPHSAVFNENKSTRTROTQVLGLTQTCETLKLQLPNCIEDPVSVILRLNF 720

DB 677 LDALDPERLSPRATFOETKNRSLSRVRLGLKAKHCENFNLLPSCVEDSVTPITLRLNF 736
QY 721 SLVGTPLSAFCNLRPLVLAEDAQRLLTALFPPEKNCENNDIQQDDLSITFSFMSLDCLVVG 780
DB 737 TLVGKPLAFENLAPMLAALAQRYFTASLPPEKNCADHICQDNLGIGSFGLKSLVVG 796
QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPPLDLISYKSVSTLQNRQSRWRLACESASSTEV 840
DB 797 SNLEINAEVWVWVNDGSDSYRTQVTFPPPLDLISYKSVSTLQNRQSRWRLACESASSTEV 854
QY 841 SGALKSTSCSNHPIFFPENSVDFTNITPDVDSKASLGNKILLKANVTSENNMPTNKTEF 900
DB 855 SQGTWSTSCRNHILFRGAQITFLATFDVSPKAVLGRLLLTANVSENNTPRTSKITF 914
QY 901 QLELPVKYAVVWVTVSHGVSTKYLNFTAS-ENTSRVMOHOYOVSNLQORSPLISLVFLVP 959
DB 915 QLELPVKYAVVWVTVSHGVSTKYLNFTAS-ENTSRVMOHOYOVSNLQORSPLISLVFLVP 974
QY 960 VRLNQTIVLDRPQVTFSENLSTSTCHTKERLPSHSDFLAELKAPVWVNGSIAVCORIQCDI 1019
DB 975 VELNOEAVWMDVEVSHPNQPSRCSSEKIAAPPASDFLAHIOKNPVLDCSIAGCLAFRCDV 1034
QY 1020 PFFGTQEEFNATLKENLSPDFWYIKTSHNHLIIVSTAELFENDSVFTLLPGQAFVRSOTE 1079
DB 1035 PSFSVQEEELDTLKENLSPDFWYIKTSHNHLIIVSTAELFENDSVFTLLPGQAFVRSOTE 1094
QY 1080 TKVEPPEVNPDLPIVGSVCGILLIALLITLALYKLGFEKQYKDMMS 1128
DB 1095 TVLEKYKHNPPLIVGSSIGLELLALITAVLYKVGFFKQYKEMME 1143

RESULT 12

US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3436; DB 9; Length 1163;
Best Local Similarity 60.6%; Pred. No. 6e-308;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGSGRVRVVGAPQEIIVAAHQSGSLYQCYDSTGSCBPI 60
DB 20 FNLDTENAMTFOENARGFGQSVVQLGSGRVRVVGAPQEIIVAAHQSGSLYQCYDSTGSCBPI 79

QY 61 RLOVPVEAVNMVSLGSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 120
DB 80 GLQVPEAVNMVSLGSLAATTSPPOLLACGPTVHOTCSENYVKGICELFGSNLRQPOK 137

QY 121 FPEALRGCPQEDSDIAFLDGSISIIIPHDFRMEKWSVTVMEQLKSKTLFSLMOYSEBF 180

Db	138	LPVSRQCBQBOQIVLIDGSSISRNFEATMNFVRAVISOFQRPSTQFSLMOFSNKF	197
Qy	181	RIHFTFKFQNNPNRSLIKPIITOLLGRTHETATGLRWVRELFINTEGAKNAFKILFLL	240
Db	198	QTHFTFEEFRSTNPLSLLASVHOLQCTYTATAIQNVVHRLPHASYGARRDAIKILIVI	257
Qy	241	TDEKFGDPIGYEDVIBELDRGVIRVIVGODAFRSEKSRQBELNIVASFPDRDHVQIN	300
Db	258	TDGKKEGSDYKDVIPMAADAARIYAIGVGLAFONRNSWKEMLDASKEPSHEHFKVE	317
Qy	301	NFEALTIQIOLREKIPALBGTOTGSSSSFEHMSQEGFAATISNGPLSLTVGSYDWAQ	360
Db	318	DFDALKDQIOLREKIPALBGTETISSSSFELEWAQEGFAVTFDQPVLGAVGSPWGS	377
Qy	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRVRQSVLGVAPRYCHIGLVAMFR	420
Db	378	GAFLYPNNMSPTFINMGOENVDEKSDYLGYSSTELALWKGVQSVLGVAPRYCHIGKAVIFI	433
Qy	421	QNTGMESNANVKTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGQVSCPL	480
Db	438	QVSRQMRKKAIEVIGTQIGSTFGASLCSVDVDTGSTDVLIGAPHYYEOTRGGQVSCPL	497
Qy	481	PRGORARWQCDAVLYGEGQGFWRFGAALTVLGDVNGDKLTDVAIGAPEGBEDNRGAYLYF	540
Db	498	PRGWR-RWWCDAVLYGEGQHPGRFGAALTVLGDVNGDKLTDVVIGAPGEENRGAVLYF	556
Qy	541	HGTSGSGISPSHSORIAGSKLSPLOYFGQSLSGODLTMDGLVDLTVGQGHVILLRSQ	600
Db	557	HGVLGPSISPSHSORIAGSKLSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR	616
Qy	601	PVLRVKAIMFNPREVARNYPECNDQVVMGKEAGEVVRVCLHVQKSTRDRLEGIOQSVWT	660
Db	617	PVLVGVSMQFPAEIPRSAPECHQVVSQGLVQSNICLYIDKRSKNLCSRDQSSVT	676
Qy	661	YDLALDSRPHSRVAFNETKXSTRQTVLGLTQCTETLKQLPNCYEDPVSPIVLRINF	720
Db	677	LDLALAPGLSPRAIFQETKNRSLSRVRVLCGLKAKCENFNILLIPSCVEDSVIPIILRNF	736
Qy	721	SLVCTPLSACGNLBPVLAEOARLFTALPFPEKNGDNLCODDLSITTFSPMSLDCLVVG	780
Db	737	TLVGKPLFLAFNLKRPMDALAQRVFTLASLPFKKNGADHICODNLGHSIFPFGKLSLNG	796
Qy	781	GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV	840
Db	797	SNLELNAEVMWINDGDSYGTITTFSHPAGLSYRYVAEGQKQQLRSLHLC--CSAPVG	854
Qy	841	SGALKSTSCSINHPIFPENSEVNTIPEVDVDSKASLGKLLKANKYTSXNMWPKNTKEP	900
Db	855	SQGTWSTSCRINELIPRGAQITFLATFVDSPPKAVGLDRLLLIANVSSNNIPETSKTIP	914
Qy	901	QLELPKYAVVMVYTSHGVSFKYLNFTAS-ENTSRVMQHOVQVSNLQORSIPISLVLVP	959
Db	915	QLELPKYAVIVVYSSHEQTKYLNFSSEKESHVAMHRQVNNLQORLPVSNINWVP	974
Qy	960	YRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAELKAPVNCSTAVCQRIODI	1019
Db	975	VELNQERAVMDVEVSHFPQNTSLRCSSEKIPAPASDFLAHQKPNVLPDSCSIAGCLRPFCDV	1034
Qy	1020	PFQIGQEPNATLKNLSFDWVYIKTSHNHLILVSTAILFNDVSVFTLLPQCAFVRSQTE	1079
Db	1035	PSFSVQELDFTLKNLSFGWVRQILQKQSVSVSAEIIIFDTSVYSQIPQGEAPWRAQTI	1094
Qy	1080	TKVEPFVFNPLPLIVGSSVGLLLALITAAALKYGLQFFKQYKDMMSSE	1128
Db	1095	TVLEKYVHNPPIPIVSGSIGILLALITAVLYKGVGFFKQYKEMMESE	1143

RESULT 13

RESULT 13
IIS-09-897-943-4

US-05-851-343-4
: Sequence 4. Application US/09891943

; sequence #; Application No. US20030077278A1

: GENERAL INFORMATION:

721 SLVGTPLSFGNLPVLAEDAORLFTALFPFKKNCNDNICODDLSTFFSMSLDCLVVG 780
 734 SLVEEPIFSPONLRFVLAAGSODLFTASLPFKKNCQDGLCEGLVTLSPSGLOTTVG 793
 781 GPREFNVTVTRNDGDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLLACESASSTEV 840
 794 SSLELANVTVWNAAGDSYGTVVSLLYPAGLSHRRVSGAQKQPHQSALELACETV-PTED 852
 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
 853 EG-LRSSRCVNHPIFHEGSGNGTIFVTVDYSYKATLGRMLMRASSENKASSSKATP 911
 901 QLELPVKYAVTVVTVSHGVSTKYLNF-TASENTSRRVMQHOYQVSNLQORSIFISLVFLVP 959
 912 QLELPVKYAVTVVTVSHGVSTKYLNF-TASENTSRRVMQHOYQVSNLQORSIFISLVFLVP 971
 960 VRLNQTWVDRPQVTFPSENLSSTCHTKERLPSHSDPLABLAKAPVNCSTAVCORIQCDI 1019
 972 VLLNGVAVWVWMEAPSQSL--PCVSEKTPQHSDFLTQISRSFMDCSADCLQFRCDV 1029
 1020 PFFGIGQSEFNATLKENLSFDWYIKTSHHLLIVGTABILFNDISVFTLLPGQCAFVRSQTE 1079
 1030 PFSVQSELDFTLKENLSFGWVRSTLQKVLVWSVAITFDTSVYSQLPGQEAFFEAQME 1089
 1080 TKVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSE 1128
 1090 MVLBEDEVYNAIPIIMGSSVGLLLALITATLYKLGFFKQYKDMSE 1138

Search completed: June 7, 2004, 17:38:47
 Job time : 41.8905 secs

APPLICANT: Van der Vieren, Monica
 TITLE OF INVENTION: No. US2003007728A1e1 Human 2
 FILE REFERENCE: 27866/35004
 CURRENT APPLICATION NUMBER: US/09/891,943
 CURRENT FILING DATE: 2001-06-26
 PRIOR APPLICATION NUMBER: 09/193,043
 PRIOR FILING DATE: 1998-11-16
 PRIOR APPLICATION NUMBER: 08/286,889
 PRIOR FILING DATE: 1994-08-05
 PRIOR APPLICATION NUMBER: 08/362,652
 PRIOR FILING DATE: 1994-12-21
 PRIOR APPLICATION NUMBER: 08/943,363
 PRIOR FILING DATE: 1997-10-03
 NUMBER OF SEQ ID NOS: 114
 SOFTWARE: Patent in Ver. 2.0
 SEQ ID NO 2
 LENGTH: 1161
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-891-943-2

Query Match 57.8%; Score 3401; DB 10; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 1.1e-304;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

1 FNLDTENAMTFORNARGQSGVQLGSRVVGAPQEIIVAAANORGLSYOCYSTGSCBPI 60
 17 FNLDVEEPTIFQDAGGQSGVQVFGSLRVVGALEVVAAVAAQTGELDYCAAAATGKQPI 76
 61 RLQVPVAVNMSLGLSAAITPPQLACGTPVHQCSTENTYVKGCLFLFGSNLQKQPK 120
 77 PLHIREAVNMSLGLTAASTNGSLRLLAGCGTLHRCVGENSYSGKSLLLGSRW-EIIQT 135
 121 FPEALRGCEQEDSIAFLDGGSIIPHPFRMKBWVSTVMEQLKSKTLPISLMOYSEEF 180
 136 VPATPECHQEMDIIVFLDGGSIQDQNFQNMKGFAVQAVMGQFEGTDTFLALMQYSNLL 195
 181 RIHFTKEFQNNPNRSLKIPTQLLGRTHRTATGLKRVVRELFTNGARKNAFKILFL 240
 196 KIHFTTQFRTSPSQSLVDPIVQLKGLTFTATGILTVTVTLFHHKNGARKGAKKILVI 255
 241 TDGEKGDPLGYEDVPELDREGVIRVYGVGDAPFSEKSRQELNVTASKPRDRHVQIN 300
 256 TDGQKYKDPLEYSVDVPOAEKAGIIRIYAGVGHAFQGPATROBELNTISAPPQDHFVKVD 315
 301 NFALKTIQNLREKIFAEIGTGTGSSSFHEMSQEGFSAALTSNGPLLLSTVGSYDMAG 360
 316 NFAALGSIQKLOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 375
 361 GVFLYTSKEKSTPINTRVDSMDNDAYLGAAAIILBNRVQSLVLCAPRYOHIGLVAMFR 420
 376 GAFLYPPNMSPTINKSQENVDNRDYLGYSTELALWQVNLVLCAPRYOHTGKAVIF 435
 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIAGAPHYETQTRGGQVSVCP 480
 436 QVSRQWRKRAEVTVGLIGSYFGASLCSVDVDSNGSTDLVLIAGAPHYETQTRGGQVSVCP 495
 481 PRQGRARWQCDVLYGEGQGPWRGFGAALTVDLVNGDKLTDVAIGAPGEENRGAVYLF 540
 496 PRQGRVQWQCDVLRGEGHPWRGFGAALTVDLVNEDKLLIDVAIGAPGEENRGAVYLF 555
 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGCAQGHVLLRSQ 600
 556 HGASESGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGCAQGHVLLRSQ 615
 601 PVLRVKAIMEFNPREVARNVFECDNVVKGKAGEVRVCLVHVKQSTRDLRREGQIOSVVT 660
 616 PVLVGVVAMEFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 661 YDLALDSGRPHSAVNETKNSRTQTVLGLTQCTETKQLPNCIEDPVSPVILRLNF 720
 674 FDLALDPGRLLTSRAIFNETKNPTLTKRKTGLGHCETKLLLLPDCVDVWVSPILHLNF 733

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds

(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKROYKDWMSBGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents 2A.*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pap:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pap:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pap:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pap:*
- 5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pap:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1	US-08-173-497-3
2	5852	99.5	1153	1	US-08-286-889-3
3	5852	99.5	1153	1	US-08-485-618-3
4	5852	99.5	1153	1	US-08-362-652-3
5	5852	99.5	1153	2	US-08-605-672-3
6	5852	99.5	1153	2	US-08-482-293A-3
7	5852	99.5	1153	2	US-08-943-363-3
8	5852	99.5	1153	3	US-09-193-043-3
9	5852	99.5	1153	4	US-09-688-307A-3
10	5852	99.5	1153	4	US-09-350-259-3
11	5821.5	98.9	1152	2	US-08-476-062A-43
12	5821.5	98.9	1152	2	PCT-US96-01314-43
13	5821.5	98.9	1152	6	5424399-2
14	3459	58.8	1163	2	US-08-476-062A-44
15	3459	58.8	1163	5	PCT-US96-01314-44
16	3436	58.4	1163	1	US-08-173-497-4
17	3436	58.4	1163	1	US-08-286-889-4
18	3436	58.4	1163	1	US-08-485-618-4
19	3436	58.4	1163	1	US-08-362-652-4
20	3436	58.4	1163	2	US-08-605-672-4
21	3436	58.4	1163	2	US-08-482-293A-4
22	3436	58.4	1163	2	US-08-943-363-4
23	3436	58.4	1163	3	US-09-193-043-4
24	3436	58.4	1163	4	US-09-688-307A-4
25	3436	58.4	1163	4	US-09-350-259-4
26	3401	57.8	1161	1	US-08-173-497-2
27	3401	57.8	1161	1	US-08-286-889-2

28	3401	57.8	1161	1	US-08-485-618-2
29	3401	57.8	1161	1	US-08-362-652-2
30	3401	57.8	1161	2	US-08-605-672-2
31	3401	57.8	1161	2	US-08-482-293A-2
32	3401	57.8	1161	2	US-08-943-363-2
33	3401	57.8	1161	3	US-09-193-043-2
34	3401	57.8	1161	4	US-09-688-307A-2
35	3401	57.8	1161	4	US-09-350-259-2
36	3385.5	57.5	1161	1	US-08-485-618-99
37	3385.5	57.5	1161	2	US-08-605-672-99
38	3385.5	57.5	1161	2	US-08-482-293A-99
39	3385.5	57.5	1161	2	US-08-943-363-99
40	3385.5	57.5	1161	3	US-09-193-043-99
41	3385.5	57.5	1161	4	US-09-688-307A-99
42	3385.5	57.5	1161	4	US-09-350-259-99
43	3224.5	54.8	1161	3	US-09-193-043-55
44	3224.5	54.8	1161	4	US-09-688-307A-55
45	3224.5	54.8	1161	4	US-09-350-259-55

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQETVAANQSGSLYQCYSTGSCBPI	60
DB	17	FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQETVAANQSGSLYQCYSTGSCBPI	76

QY 61 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQOPK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQOPK 136
QY 121 FPALRGCCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEOLKSKTFLPSLMQYSEEP 180
DB 137 FPALRGCCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEOLKSKTFLPSLMQYSEEP 196
QY 181 RIHFTFKFQNNPNRSLKPIITQLLGRTHATGLRKVVRELNTNGARKNAKILFL 240
DB 197 RIHFTFKFQNNPNRSLKPIITQLLGRTHATGLRKVVRELNTNGARKNAKILFL 256
QY 241 TDCEKFGDPLGYEDVITPEADREGVIRVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDCEKFGDPLGYEDVITPEADREGVIRVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 316
QY 361 GVELYTSKSKSTINTRVDSNDAYLGVAAIILNRVQSLVGLGAPVQIHGLVAMER 420
DB 377 GVELYTSKSKSTINTRVDSNDAYLGVAAIILNRVQSLVGLGAPVQIHGLVAMER 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEQTRGQVSVCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEQTRGQVSVCP 496
QY 481 PRGRARWQCDVLYGEOGQWPGFAGALTVLGVNDKLTVDVAIGARGEEDNRGAVLF 540
DB 497 PRGRARWQCDVLYGEOGQWPGFAGALTVLGVNDKLTVDVAIGARGEEDNRGAVLF 556
QY 541 HGTSGSGISFSHSORIASKSLSPRLQYFGQSLGGQDLTMDGLVLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISFSHSORIASKSLSPRLQYFGQSLGGQDLTMDGLVLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFPNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVT 660
DB 617 PVLRVKAIMFPNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQCTETLKLQPCNIEDPVSIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQCTETLKLQPCNIEDPVSIVLRNF 736
QY 721 SLVGTPLSAGNLRPLVLAEDAQLFTALPFPKXNCGNDNICQDDLSITFSFMSLDCILV 780
DB 737 SLVGTPLSAGNLRPLVLAEDAQLFTALPFPKXNCGNDNICQDDLSITFSFMSLDCILV 796
QY 781 GPREFNVTIVRNDGEDSYTQVTFPPDLISVRKYSTLQNSORSWRLACESASSTEV 840
DB 797 GPREFNVTIVRNDGEDSYTQVTFPPDLISVRKYSTLQNSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 900
DB 857 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 916
QY 901 QLELPVYAVYVMTVSHGVSTKYNFTASENTSRVMOHQVQVNLGORSIPISLVLV 960
DB 917 QLELPVYAVYVMTVSHGVSTKYNFTASENTSRVMOHQVQVNLGORSIPISLVLV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCQRIQDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCQRIQDIP 1036
QY 1021 PFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQAFVRSQTET 1080
DB 1037 PFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPGQAFVRSQTET 1096
QY 1081 KVPEFVFNPLIPLVSGSVGLLALITAALYKLGFFKQYKDMWSEGGPPGAEQ 1137
DB 1097 KVPEFVFNPLIPLVSGSVGLLALITAALYKLGFFKQYKDMWSEGGPPGAEQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQOPK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQOPK 136
QY 121 FPALRGCCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEOLKSKTFLPSLMQYSEEP 180
DB 137 FPALRGCCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEOLKSKTFLPSLMQYSEEP 196
QY 181 RIHFTFKFQNNPNRSLKPIITQLLGRTHATGLRKVVRELNTNGARKNAKILFL 240
DB 197 RIHFTFKFQNNPNRSLKPIITQLLGRTHATGLRKVVRELNTNGARKNAKILFL 256
QY 241 TDCEKFGDPLGYEDVITPEADREGVIRVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDCEKFGDPLGYEDVITPEADREGVIRVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 316
QY 301 NFEALKTIONQLREKIPALEGTQTSSTSSFFHEMSQEGFSAAITSNGLPSTVGSYDWA 360
DB 317 NFEALKTIONQLREKIPALEGTQTSSTSSFFHEMSQEGFSAAITSNGLPSTVGSYDWA 376

361 GVFLYTSKEKSTFIINRTVDSMDNDAYLGAAAILRNVRVSLVGLGAPRYOHIGLVAMFR 420
377 GVFLYTSKEKSTFIINRTVDSMDNDAYLGAAAILRNVRVSLVGLGAPRYOHIGLVAMFR 436
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSVCP 480
437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSVCP 496
481 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVGVNGDKLTDAVAGAPBEDNRGAVYLF 540
497 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVGVNGDKLTDAVAGAPBEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDTVGAQSHVLLRSQ 600
557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDTVGAQSHVLLRSQ 616
601 PVLRYKATMEFNPREVARNVFCNQVVKGEKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLRYKATMEFNPREVARNVFCNQVVKGEKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGRPHSRVFNETKSTRQTVGLGTQTCETLKLQPLNCIEDPVPVILRLNF 720
677 YDLALDSGRPHSRVFNETKSTRQTVGLGTQTCETLKLQPLNCIEDPVPVILRLNF 736
721 SLVGTPLSAFNLRPVLAEDAQLFTALPPFPKNCNDNICODDLSITFSMSLDCLVVG 780
737 SLVGTPLSAFNLRPVLAEDAQLFTALPPFPKNCNDNICODDLSITFSMSLDCLVVG 796
781 GRPREFNVTVTVNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
797 GRPREFNVTVTVNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKVTSNNPRNTKTEF 900
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANKVTSNNPRNTKTEF 916
901 QLELPVKYAVVWVTSHGYSTKYLNFATASNTSRVMQHOYQVSNLQORSPLTSVLPLVPV 960
917 QLELPVKYAVVWVTSHGYSTKYLNFATASNTSRVMQHOYQVSNLQORSPLTSVLPLVPV 976
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1036
1021 FPGIOBEFNATLGNLSFDWYIKTSHNLLIYSTAEILFENDSVFTLLPQGAFFVRSQTET 1080
1037 FPGIOBEFNATLGNLSFDWYIKTSHNLLIYSTAEILFENDSVFTLLPQGAFFVRSQTET 1096
1081 KVEPFEVPNPPLTVGSSVGGLLLLALITAAALYKLGFFKQYKQYKQYKQYKQYKQYKQY 1137
1097 KVEPFEVPNPPLTVGSSVGGLLLLALITAAALYKLGFFKQYKQYKQYKQYKQYKQYKQY 1153

RESULT 3
US-08-485-618-3
Sequence 3, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTPOENARFGQSVVQLQSGRVVVGAPQEIIVAAOQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTPOENARFGQSVVQLQSGRVVVGAPQEIIVAAOQSGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNKSIGLSLAATTSPQLLACGPTVHOTCSENTYVVKGLCLFGLSGLNRQOQOK 120
DB 77 RLQVPVEAVNKSIGLSLAATTSPQLLACGPTVHOTCSENTYVVKGLCLFGLSGLNRQOQOK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEFVSTVMEQLKSKTFLSLAQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEFVSTVMEQLKSKTFLSLAQYSEEF 196
QY 181 RIHFTFKPQNNPNSRLIKETITOLLGRTHATGLRVVRELENTNGARKNAFKILFLL 240
DB 197 RIHFTFKPQNNPNSRLIKETITOLLGRTHATGLRVVRELENTNGARKNAFKILFLL 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPRSEKSRQELMTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPRSEKSRQELMTVASKPPRDHVFQIN 316
QY 301 NPEALKTIQNLREKIPIAETGOTGSSSSFEHMSQEGFSAAITNSGILLSTVGSYDWAG 360
DB 317 NPEALKTIQNLREKIPIAETGOTGSSSSFEHMSQEGFSAAITNSGILLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFIINRTVDSMDNDAYLGAAAILRNVRVSLVGLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFIINRTVDSMDNDAYLGAAAILRNVRVSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSVCP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVGVNGDKLTDAVAGAPBEDNRGAVYLF 540
DB 497 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVGVNGDKLTDAVAGAPBEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDTVGAQSHVLLRSQ 600


```
Db 557 HGTSGSISPSHSORISAGSKLSPRLQYFGQSLGQQLTWDGLVDTLVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 840
Db 797 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 856
Qy 841 SGALKSTSCSINIPFPENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINIPFPENSEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPVKYAVVMVTSKGVTKYLINFASNTSRVQHQYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSKGVTKYLINFASNTSRVQHQYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNQTVLWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNCSTAVCORIOCDIP 1020
Db 977 RLNQTVLWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNCSTAVCORIOCDIP 1036
Qy 1021 PFGIQEBFNATLGNLSDFYIKTSHNHLIVSTAEILFNDSTFTLLPGQGAFFVRSQTEF 1080
Db 1037 PFGIQEBFNATLGNLSDFYIKTSHNHLIVSTAEILFNDSTFTLLPGQGAFFVRSQTEF 1096
Qy 1081 KVEPPEVNPPLIVGSSVGLLILALITAALYKLGFFKQYKDMSEGGPPGABEQ 1137
Db 1097 KVEPPEVNPPLIVGSSVGLLILALITAALYKLGFFKQYKDMSEGGPPGABEQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5786850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
```

```
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 278666/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3
```

Query Match 99.5%; Score 5852; DB 1; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRQSLQCDYSTGSCERI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRQSLQCDYSTGSCERI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFPGNLNQPOPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFPGNLNQPOPK 136
Qy 121 FPEARLGCPOEDSDAFILIDSGSIIIPHDPRMKEMVSTVMBOLKSKTFLSLMOYSBEF 180
Db 137 FPEARLGCPOEDSDAFILIDSGSIIIPHDPRMKGFVSTVMEQLKSKTFLSLMOYSBEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFNITNGARKNAKILFVL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFNITNGARKNAKILFVL 256
Qy 241 TDGEKGGPLGVEDVIPILDRREGVIRYVGVGDAPRSKRSQELNNTVASKPRDHVFOIN 300
Db 257 TDGEKGGPLGVEDVIPILDRREGVIRYVGVGDAPRSKRSQELNNTVASKPRDHVFOIN 316
Qy 301 NFEALKTIQNLREKIPAIETGTQTSSSSFEHMSQEGFSAAITSNGPLLTSGSYDMAQ 360
Db 317 NFEALKTIQNLREKIPAIETGTQTSSSSFEHMSQEGFSAAITSNGPLLTSGSYDMAQ 376
Qy 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEBQTRGGQSVCPPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEBQTRGGQSVCPPL 496
Qy 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPSEHNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPSEHNRGAVYLF 556
Qy 541 HGTSGSISPSHSORISAGSKLSPRLQYFGQSLGQQLTWDGLVDTLVGAQGHVLLRSQ 600
Db 557 HGTSGSISPSHSORISAGSKLSPRLQYFGQSLGQQLTWDGLVDTLVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTRDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQSRWRLACSSASTEV 840
```

Db 797 GPREFNVTVVRNDGDSYKQTWTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENTITPDVDSKASLGNKLLKXVYTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTENTITPDVDSKASLGNKLLKXVYTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVWVTSYHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLV 960
Db 917 QLELPVKYAVYVWVTSYHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLV 976
Qy 961 RLNOTVIMDRPQVTPSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTPSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1036
Qy 1021 FFGIOBEFNATLGNLSFDWYIKTNSHLLIVSTAEILEFNDVSFTLLPQOGAFVRSQTET 1080
Db 1037 FFGIOBEFNATLGNLSFDWYIKTNSHLLIVSTAEILEFNDVSFTLLPQOGAFVRSQTET 1096
Qy 1081 KVEPEVNPPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSBGGPPGASBPQ 1137
Db 1097 KVEPEVNPPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSBGGPPGASBPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,899
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-605-672-3

Query Match
Best Local Similarity 99.5%; Score 5852; DB 2; Length 1153;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLTENAMTPQENARFGQSVVQLQGRVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLTENAMTPQENARFGQSVVQLQGRVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATPPQLLACGPTVHTQCTSENTYVKGCLCFGLFQSNLRQOQPK 120
Db 77 RLQVPVAVNMSLGLSLAATPPQLLACGPTVHTQCTSENTYVKGCLCFGLFQSNLRQOQPK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEWSTVMBOLKSKTSLSLMOYSSEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQKCKSKTSLSLMOYSSEF 196
Qy 181 RIHFTFEFQNNPNSRIKPIITOLLGRTHATGLRKVVRELENTNCARKNAPKILPUL 240
Db 197 RIHFTFEFQNNPNSRLVKPIITOLLGRTHATGLRKVVRELENTNCARKNAPKILVVI 256
Qy 241 TDGEKFGDPLGYEDVIPELDREGVIRVYVIGDAPFRSEKSRQELNTVASKPRDHVFOIN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRVYVIGDAPFRSEKSRQELNTIASKPRDHVFOIN 316
Qy 301 NFEALKTIONOLREKIPIAIEGTOTGSSSPHEHNSOEGFSAITNSGELLSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKIPIAIEGTOTGSSSPHEHNSOEGFSAITNSGELLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNSTDLVILGAPHYHQTGQGVSVLCP 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNSTDLVILGAPHYHQTGQGVSVLCP 496
Qy 481 PRGQARWQCDAYLVGEQGPWGRFGAALTIVLGVNGDKLTDAVAGPEEDNRGAVILF 540
Db 497 PRGQARWQCDAYLVGEQGPWGRFGAALTIVLGVNGDKLTDAVAGPEEDNRGAVILF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVBLTVGAQHVLLLRSQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVBLTVGAQHVLLLRSQ 616
Qy 601 PVLRVKATMEFNPREVARNVFECDNDQVVKGEAGSVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKATMEFNPREVARNVFECDNDQVVKGEAGSVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGNLRVLAEDAQRLEFALPPEKKGNDNICODDLSITPFSMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRVLAEDAQRLEFALPPEKKGNDNICODDLSITPFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYKQTWTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYKQTWTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENTITPDVDSKASLGNKLLKXVYTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTENTITPDVDSKASLGNKLLKXVYTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVWVTSYHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLV 960
Db 917 QLELPVKYAVYVWVTSYHGVSTKYLANFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLV 976
Qy 961 RLNOTVIMDRPQVTPSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTPSENLSSCTKTERLPSSDPLAELRKAPVNCSTAVCORIQCDIP 1036

QY 1021 FFGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGGAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGGAFVRSOTET 1096
QY 1081 KVEPFVEVPNPLPLIVGSSVGGILLALITAALYKLGFFKROKMDMSEGGPPGABPQ 1137
DB 1097 KVEPFVEVPNPLPLIVGSSVGGILLALITAALYKLGFFKROKMDMSEGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-5402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-293A-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFGQSVVQLOGSRVVVCGAPQBIYAANORGSLYQCDYSTGSCBPI 60
DB 17 ENLDTENAMTFQENARGFGQSVVQLOGSRVVVCGAPQBIYAANORGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGICFLPGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGICFLPGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKVFSTWMEQLKSKTFLSLMQYSEEP 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKVFSTWMEQLKSKTFLSLMQYSEEP 196
QY 181 RIHFTFEKFFQNPAPRSLSIKPITQILGRTHATGLRKVKVRELFNITNGARQAKILPELL 240
DB 197 RIHFTFEKFFQNPAPRSLSIKPITQILGRTHATGLRKVKVRELFNITNGARQAKILVVI 256
QY 241 TDGKFGDPLGYEDVPELDRGVIRVYVIGVDAPRSEKSRQELNTVASKPDRPHVFOIN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRVYVIGVDAPRSEKSFQELNTIASKPDRPHVFOIN 316
QY 301 NPEALKTIQNLREKIPAIETOTGSSSSPHEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NPEALKTIQNLREKIPAIETOTGSSSSPHEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVELYTSKSKSTFINMTREVDSDMDADVLGYAAAILNRNVQSLVGLGAPRYOHIGLVAMER 420
DB 377 GVELYTSKSKSTFINMTREVDSDMDADVLGYAAAILNRNVQSLVGLGAPRYOHIGLVAMER 436
QY 421 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 480
DB 437 QNTGWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 496
QY 481 PRGQARWQCDVLYGEGQDPWGRFGALTVLGDVNGDKLTDVAGAPGEDNRCVLYF 540
DB 497 PRGQARWQCDVLYGEGQDPWGRFGALTVLGDVNGDKLTDVAGAPGEDNRCVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQHVLALLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQHVLALLRSQ 616
QY 601 PVLRVKALMEFNPREVARNFECDQVYVKGEGEVRVCLHVOKSTRDLREGQOSVVT 660
DB 617 PVLRVKALMEFNPREVARNFECDQVYVKGEGEVRVCLHVOKSTRDLREGQOSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKSTTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAPGNLAPVLAEDAQRLLFTALPPFEKKGNDNICQDDLSITFSFMSLDCLVNG 780
DB 737 SLVGTPLSAPGNLAPVLAEDAQRLLFTALPPFEKKGNDNICQDDLSITFSFMSLDCLVNG 796
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENGSEVTVNITPDVDSKASLGNKLLKANVTSENNMPRNTKTEP 900
DB 857 SGALKSTSCSINHPIPPENGSEVTVNITPDVDSKASLGNKLLKANVTSENNMPRNTKTEP 916
QY 901 QLELPVKYAVVMVYTSKGVSTKYLNTASNTSRVMQHQYQVSNLQGRSLPISLVFLVFP 960
DB 917 QLELPVKYAVVMVYTSKGVSTKYLNTASNTSRVMQHQYQVSNLQGRSLPISLVFLVFP 976
QY 961 RLNQTVIWDPRPQVTPSENLSTCHTKERLPSHSDFLABLRKAPVWNCISVAVCORIQCDIP 1020
DB 977 RLNQTVIWDPRPQVTPSENLSTCHTKERLPSHSDFLABLRKAPVWNCISVAVCORIQCDIP 1036
QY 1021 FPGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGGAFVRSOTET 1080
DB 1037 FPGIOEFNATLKGNSLSPWYIKTSHNLLIYSTABILENDVSFTLLPGGAFVRSOTET 1096
QY 1081 KVEPFVEVPNPLPLIVGSSVGGILLALITAALYKLGFFKROKMDMSEGGPPGABPQ 1137
DB 1097 KVEPFVEVPNPLPLIVGSSVGGILLALITAALYKLGFFKROKMDMSEGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3956
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFCQSVVVLQGSRVVVGAPQEIIVAAANQSGLYOCDSYTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFCQSVVVLQGSRVVVGAPQEIIVAAANQSGLYOCDSYTGSCBPI 76

QY 61 RLQVPVEAVNMVSLGLSLAATTGPPQLLACGPTVHOTCSENTYVKGCLPLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMVSLGLSLAATTGPPQLLACGPTVHOTCSENTYVKGCLPLFGSNLRQOPQK 136

QY 121 FEALRGCEQSDIAFLIDGSGSIIPHPFRMKWVSTVMEQLKSKTLFSLMOYSEFP 180
DB 137 FEALRGCEQSDIAFLIDGSGSIIPHPFRMKWVSTVMEQLKSKTLFSLMOYSEFP 196

QY 181 RHFTFKFQNNPNSLKIPIQLGRTHATGLRKVVRELFNITNGARKNAKILFL 240
DB 197 RHFTFKFQNNPNSLKIPIQLGRTHATGLRKVVRELFNITNGARKNAKILFL 256

QY 241 TDGEKFGDPLGYEDVITPELDREGVIRYIVGVGDAPFSEKSRQELNVTASKPPRDVDFQIN 300
DB 257 TDGEKFGDPLGYEDVITPELDREGVIRYIVGVGDAPFSEKSRQELNVTASKPPRDVDFQIN 316

QY 301 NFEALATIOQREKIPALEGTOTGSSSFHEMSQEGFSAATISNGPLLSVTGSDVWAG 360
DB 317 NFEALATIOQREKIPALEGTOTGSSSFHEMSQEGFSAATISNGPLLSVTGSDVWAG 376

QY 361 GVFLYTSKEKSTFTINMTREVDSXNDAYLGAAAAIILNRVQSLVLTGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFTINMTREVDSXNDAYLGAAAAIILNRVQSLVLTGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKCTQICAGYFAGSILCSVDVDSNGSTDLVLICAGPHYYPQTRGGQSVVCP 480
DB 437 QNTGMWESNANVKCTQICAGYFAGSILCSVDVDSNGSTDLVLICAGPHYYPQTRGGQSVVCP 496

QY 481 PRGQARWQCDVLYGBOGQPNWGRFGAALTVDLVGNGDKLTDVAITGAPGEDNRCAGVTLF 540
DB 497 PRGQARWQCDVLYGBOGQPNWGRFGAALTVDLVGNGDKLTDVAITGAPGEDNRCAGVTLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPLRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLARSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLARSQ 616

QY 601 PVLAVKALMEENPREVARNVFECNDQVVKGEAGEVRVCLAVKSTRDRLEGGIQSVVT 660
DB 617 PVLAVKALMEENPREVARNVFECNDQVVKGEAGEVRVCLAVKSTRDRLEGGIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKSTTRRQTVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
DB 677 YDLALDSGRPHSRVAFNETKSTTRRQTVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736

QY 721 SLVGTPLSAPGNLRPVLAEQAQRLFTALPPPEKNCNDNICODDLSTITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEQAQRLFTALPPPEKNCNDNICODDLSTITFSFMSLDCLVVG 796

QY 781 GPRFENVTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPRFENVTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTNITPFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITPFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916

QY 901 QLELPKYAVTVMTVSHGVSTKYLNFTASENTSRVMOHQYOVSNLQORSPLSLVFLVPV 960
DB 917 QLELPKYAVTVMTVSHGVSTKYLNFTASENTSRVMOHQYOVSNLQORSPLSLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCIAVCQRIQCDDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCIAVCQRIQCDDIP 1036

QY 1021 PFGQREBNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVPTLLPQCGAFVRSOTET 1080
DB 1037 PFGQREBNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVPTLLPQCGAFVRSOTET 1096

QY 1081 KVEPPEVENPPLIVGSSVGGILLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPPQ 1137
DB 1097 KVEPPEVENPPLIVGSSVGGILLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARFGQSVVQLGSGVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGQSVVQLGSGVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSGLSLAATTSPPQLLAGCTVHQTCSNTYVKGECFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATTSPPQLLAGCTVHQTCSNTYVKGECFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEVSTVMEQLKSKTLPFLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEVSTVMEQLKSKTLPFLMOYSEEF 196

Qy 181 RIHFTKEFQNNPNRSLIKPITOLHGRTHATGLRKVRELFTNITNGARKNAFKILFLL 240
Db 197 RIHFTKEFQNNPNRSLIKPITOLHGRTHATGLRKVRELFTNITNGARKNAFKILFLL 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPPRDHVFQIN 316

Qy 301 NFEALKTIONQLREKI FAIEGTQTGSSSFHEMSQEGFSAATITNGPILSTVGSYDWAG 360

US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARFGQSVVQLGSGVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGQSVVQLGSGVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSGLSLAATTSPPQLLAGCTVHQTCSNTYVKGECFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATTSPPQLLAGCTVHQTCSNTYVKGECFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEVSTVMEQLKSKTLPFLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEVSTVMEQLKSKTLPFLMOYSEEF 196

Qy 181 RIHFTKEFQNNPNRSLIKPITOLHGRTHATGLRKVRELFTNITNGARKNAFKILFLL 240
Db 197 RIHFTKEFQNNPNRSLIKPITOLHGRTHATGLRKVRELFTNITNGARKNAFKILFLL 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPPRDHVFQIN 316

Qy 301 NFEALKTIONQLREKI FAIEGTQTGSSSFHEMSQEGFSAATITNGPILSTVGSYDWAG 360

Db 317 NFEALKTIONQREKI PAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAILLRNVQSIVLGPARYOHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAILLRNVQSIVLGPARYOHIGLVAMPR 436
Qy 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVVCPL 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVVCPL 496
Qy 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNKGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNKGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYEGQSISGGQDLTMDGLVDLTGACGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLSPRLQYEGQSISGGQDLTMDGLVDLTGACGHVLLRSQ 616
Qy 601 PVLRVKALMEFNPVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKALMEFNPVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNFKSTRQTOVLGHTOTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNFKSTRQTOVLGHTOTCETLKLQLPNCIEDPSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGSDSYRTQVTFPPDLDSYKRVSTLQORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGSDSYRTQVTFPPDLDSYKRVSTLQORSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKAVMTVSHGVSTKYLNFASNTSRVQOYQVSNLQORSQSWRLACESASSTEV 960
Db 917 QLELPVKAVMTVSHGVSTKYLNFASNTSRVQOYQVSNLQORSQSWRLACESASSTEV 976
Qy 961 RLNQTVIWDROPVTPSENLSSTCHTKERLPSSHDFLAELRKAPVVCISIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDROPVTPSENLSSTCHTKERLPSSHDFLAELRKAPVVCISIAVCQRIQCDIP 1036
Qy 1021 PFGIQEEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILLFNDVSFVTLPLPGQAFVRSQDET 1080
Db 1037 PFGIQEEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILLFNDVSFVTLPLPGQAFVRSQDET 1096
Qy 1081 KVEPEVENPLPLIVGSSVGGILLALITALYKLGPFKROYKDMWSEGGPPGAEPO 1137
Db 1097 KVEPEVENPLPLIVGSSVGGILLALITALYKLGPFKROYKDMWSEGGPPGAEPO 1153

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTEANMTPOENARGFGQSVVQLQGSRVVVCAPQEIIVAAQORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEANMTPOENARGFGQSVVQLQGSRVVVCAPQEIIVAAQORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTSPPOLLACQPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTSPPOLLACQPTVHQTCSNTYVKGCLFPGSNLRQOPQK 136
Qy 121 FPALRCPCOEDSDIAPLIDSGSIIIPHDPRRMKEWVSTVMEOLKSKTLPFLMOYSEEP 180
Db 137 FPALRCPCOEDSDIAPLIDSGSIIIPHDPRRMKEFVSTVMEOLKSKTLPFLMOYSEEP 196
Qy 181 RIHFTPEKQNNPNRSLIKPITQLGRHTATGLRKVRELFNITNGARKNAFKIILVL 240
Db 197 RIHFTPEKQNNPNRSLIKPITQLGRHTATGLRKVRELFNITNGARKNAFKIILVL 256
Qy 241 TDCEKPGDPIGYEDVTPELDREGVIRVYGVGDAFRSEKROELNTVASKPPRPHVQIN 300
Db 257 TDCEKPGDPIGYEDVTPELDREGVIRVYGVGDAFRSEKROELNTVASKPPRPHVQIN 316
Qy 301 NFZALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFZALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAILLRNVQSIVLGPARYOHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAILLRNVQSIVLGPARYOHIGLVAMPR 436
Qy 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVVCPL 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVVCPL 496
Qy 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNKGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNKGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYEGQSISGGQDLTMDGLVDLTGACGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLSPRLQYEGQSISGGQDLTMDGLVDLTGACGHVLLRSQ 616
Qy 601 PVLRVKALMEFNPVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKALMEFNPVARNVFECDQVVKGEAGVRVCLHVOKSTRDLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNFKSTRQTOVLGHTOTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNFKSTRQTOVLGHTOTCETLKLQLPNCIEDPSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGSDSYRTQVTFPPDLDSYKRVSTLQORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGSDSYRTQVTFPPDLDSYKRVSTLQORSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900

Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 916
Qy 901 QLELPVKYAVYVMTSHGVSSTKYNFTASENTRVMOHOYQVSNLQORSPLSLVPLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSSTKYNFTASENTRVMOHOYQVSNLQORSPLSLVPLVPV 976
Qy 961 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1020
Db 977 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQDET 1080
Db 1037 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQDET 1096
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKMMSEGGPPGABPQ 1137
Db 1097 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKMMSEGGPPGABPQ 1153

RESULT 11

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Annacut, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-476-062A-43

Query Match 98.94; Score 5821.5; DB 2; Length 1152;
Best Local Similarity 98.94; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGOSVVQLOQSRVVVGAPOIIVAAQNGSLYQCDYSTGSCSEPI 60
Db 17 FNLDTENAMTFOENARGFGOSVVQLOQSRVVVGAPOIIVAAQNGSLYQCDYSTGSCSEPI 76
Qy 61 RLQVPEAVANMSLGLSLAAATTPPOLLAGCPTVHTQCTSENTRYKGLCHLPGSNLRQOPQK 120
Db 77 RLQVPEAVANMSLGLSLAAATTPPOLLAGCPTVHTQCTSENTRYKGLCHLPGSNLRQOPQK 136
Qy 121 PPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEWSTVMEQKKSKTKLPSLMQYSSEF 180
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIIPHDFFRMKEWSTVMEQKKSKTKLPSLMQYSSEF 196
Qy 181 RIHTTFKFPQNNPNSRLIIPITOLLGRTHATGLRKVRBELFNI TNGARKNAKILPLL 240
Db 197 RIHTTFKFPQNNPNSRLIIPITOLLGRTHATGLRKVRBELFNI TNGARKNAKILPLL 256
Qy 241 TDGKFGDPLGYEDVPELDRGVRVYVIGVDAPFRSEKROELNLTVAASKPRDRHVFQIN 300
Db 257 TDGKFGDPLGYEDVPELDRGVRVYVIGVDAPFRSEKROELNLTVAASKPRDRHVFQIN 316
Qy 301 NFEALKTIQNLQRKIPFAIEGTQTGSSSSPHEHMSQSGFSAITNSNGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLQRKIPFAIEGTQTGSSSSPHEHMSQSGFSAITNSNGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPRYQHTGQGVQSVCP 480
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPRYQHTGQGVQSVCP 496
Qy 481 PRGORARWCDALYVGEQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNDRGAVLP 540
Db 497 PRG-RARWQCDALYVGEQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNDRGAVLP 555
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLGQQLTMDGLVDLTVGAQCHVLLRSQ 600
Db 556 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLGQQLTMDGLVDLTVGAQCHVLLRSQ 615
Qy 601 PVLRVKAIEMFNPREVARNVFECDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKAIEMFNPREVARNVFECDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAPGNLAPVLAEDAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAPGNLAPVLAEDAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREPNVTVVRNDEGDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWSRLACESASSTEV 840
Db 796 GPREPNVTVVRNDEGDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWSRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 915
Qy 901 QLELPVKYAVYVMTSHGVSSTKYNFTASENTRVMOHOYQVSNLQORSPLSLVPLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSSTKYNFTASENTRVMOHOYQVSNLQORSPLSLVPLVPV 975
Qy 961 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1020
Db 976 RLNTQVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDDIP 1035
Qy 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQDET 1080
Db 1036 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQDET 1095
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKROQKMMSEGGPPGABPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1152
|||||
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaut
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43
Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;
QY 1 FNLDTENAMTFOENARGCGQSVVQVQGGSRVVGAPQEIIVAAVNOGSLYCCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGCGQSVVQVQGGSRVVGAPQEIIVAAVNOGSLYCCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGTVHQTCSNTYVVKGLCFPLGSLNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGTVHQTCSNTYVVKGLCFPLGSLNLRQPOK 136
QY 121 PFEALRGCPQSDIAFLIDGSGSIIPDFRPMKWSVSTWEOQLKSKTLPSLMQYSEEP 180
Db 137 PFEALRGCPQSDIAFLIDGSGSIIPDFRPMKWSVSTWEOQLKSKTLPSLMQYSEEP 196
QY 181 RHFTPFKEFQNNPNRSLKPTQLLGRTHATGLRKVVRELFTNITNGARKNAKFLPL 240
Db 197 RHFTPFKEFQNNPNRSLKPTQLLGRTHATGLRKVVRELFTNITNGARKNAKFLPL 256
QY 241 TDGKPGDPLGYEDVPIPELDREGVIRYVIGVDGAFRSEKSRQELMTVAKSPRDRHVFOIN 300
Db 257 TDGKPGDPLGYEDVPIPELDREGVIRYVIGVDGAFRSEKSRQELMTVAKSPRDRHVFOIN 316
QY 301 NFEALKTIONLRKLEFLEGTCOTGSSSFEHMSQEGFSAITNGPILLSVTGSDYDAG 360
Db 317 NFEALKTIONLRKLEFLEGTCOTGSSSFEHMSQEGFSAITNGPILLSVTGSDYDAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAALILNRRVQSLVIGAPRYOHIGLVAMER 420
Db 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAALILNRRVQSLVIGAPRYOHIGLVAMER 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYTQRTGQGVSCPL 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYTQRTGQGVSCPL 496
QY 481 PRGORARWOCDAVLKXGOGOPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540
Db 497 PRG-RASWQCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 555
QY 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAGQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAGQGHVLLRSQ 615
QY 601 PVLRVKALMEFNPREVARNVPECNDQVYVKGKEAGEVRVCLHVOKSTDRIRREGOIQSVT 660
Db 616 PVLRVKALMEFNPREVARNVPECNDQVYVKGKEAGEVRVCLHVOKSTDRIRREGOIQSVT 675
QY 661 YDLALDSGRPHSRAPVNETKNSRRQTVGLGLTQTCETLKLQFPNCIEDPVSPIVLRINF 720
Db 676 YDLALDSGRPHSRAPVNETKNSRRQTVGLGLTQTCETLKLQFPNCIEDPVSPIVLRINF 735
QY 721 SLVGTPLSAFQNLKRPVLAEDAQRLEFALPPEKNCNGNDNICODDLSTTFSPMSLDCLVVG 780
Db 736 SLVGTPLSAFQNLKRPVLAEDAQRLEFALPPEKNCNGNDNICODDLSTTFSPMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSORSWRLLACESASSTEV 840
Db 796 GPRESNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNSORSWRLLACESASSTEV 855
QY 841 SGALKSTSCSINHPPIPPENSEVFNITFDVDSKASLQNLKLLKANVTSENNMPRTNKTET 900
Db 856 SGALKSTSCSINHPPIPPENSEVFNITFDVDSKASLQNLKLLKANVTSENNMPRTNKTET 915
QY 901 QLELPVKYAVYVTVSHGVSTKYLNTASNTSRVMOHOYQVSNLQORSPLISLVELVPV 960
Db 916 QLELPVKYAVYVTVSHGVSTKYLNTASNTSRVMOHOYQVSNLQORSPLISLVELVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKAIPVNCISIAVCQRIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKAIPVNCISIAVCQRIQCDIP 1035
QY 1021 PFGIOEFNATLKNLSFDMYIKTSHNELLIVSTAEILFNDVSPTLLPGCGAFVRSOTET 1080
Db 1036 PFGIOEFNATLKNLSFDMYIKTSHNELLIVSTAEILFNDVSPTLLPGCGAFVRSOTET 1095
QY 1081 KVSEPFVNPPLITVCGSVGGLLALLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1096 KVSEPFVNPPLITVCGSVGGLLALLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1152
RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2:
; LENGTH: 1152
5424399-2
Query Match 98.9%; Score 5821.5; DB 6; Length 1152;

Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLGSRVYVAGPOEIVAAANQSGSLYOCYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLGSRVYVAGPOEIVAAANQSGSLYOCYSTGSCBPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQCSENTYVKGCLPFLGNSLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQCSENTYVKGCLPFLGNSLRQPOK 136

Qy 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDPRRMEKQVSTWMEOLKSKTFLSLMOYSEEP 180
Db 137 FPEALRGCCQEDSDIAFLIDGSGSIIPHDPRRMEKQVSTWMEOLKSKTFLSLMOYSEEP 196

Qy 181 RIHFTFEKFNQNNPRSLKIPITQLGRTHRTATGLAKVVRLEFNTINGARKNAFKILFL 240
Db 197 RIHFTFEKFNQNNPRSLKIPITQLGRTHRTATGLAKVVRLEFNTINGARKNAFKILFL 256

Qy 241 TDGEKGDPLGYEDVPEIDREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHDVQFN 300
Db 257 TDGEKGDPLGYEDVPEIDREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHDVQFN 316

Qy 301 NFPAKTIQNLREKIFAIBGTGTGSSSPFHEMSQEGFSAITNSGPLISTVGSYDMAG 360
Db 317 NFPAKTIQNLREKIFAIBGTGTGSSSPFHEMSQEGFSAITNSGPLISTVGSYDMAG 376

Qy 361 GVFLYTSKEKSTINFRVDSNDAYLVAAAIILNRVQSLVGLGAPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTINFRVDSNDAYLVAAAIILNRVQSLVGLGAPYOHIGLVAMFR 436

Qy 421 QNTGMESNANVKGTOIGAFGASLCSVDVDSNGSDTLVLI GAPHYEYOTRGQVSVCP 480
Db 437 QNTGMESNANVKGTOIGAFGASLCSVDVDSNGSDTLVLI GAPHYEYOTRGQVSVCP 496

Qy 481 PRGORARWQCDAYLGEQGFWRPGALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 540
Db 497 PRG-RARWQCDAYLGEQGFWRPGALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 555

Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTWDGLVDTVGAQGHVLLRQ 600
Db 556 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTWDGLVDTVGAQGHVLLRQ 615

Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKQKEAGEVRVCLHVKSTRDLRREGQIQSVVT 660
Db 616 PVLVRKAIMEFNPREVARNVFCNDQVVKQKEAGEVRVCLHVKSTRDLRREGQIQSVVT 675

Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPCNIEDPVSIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPCNIEDPVSIVLRNF 735

Qy 721 SLVGTPLSAFNGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CQDDLSTITFSFMSLDCLVG 780
Db 736 SLVGTPLSAFNGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CQDDLSTITFSFMSLDCLVG 795

Qy 781 GPREFNVTVRNDGDSYQTVTFPPPLDLSVRKUSTLQNRQSRWRLACASSTEV 840
Db 796 GPREFNVTVRNDGDSYQTVTFPPPLDLSVRKUSTLQNRQSRWRLACASSTEV 855

Qy 841 SGALKSTSCSINRPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINRPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915

Qy 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQVQVNSNGORSPLISLVLFPV 960
Db 916 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQVQVNSNGORSPLISLVLFPV 975

Qy 961 RLQNTVLDWRPQVTFSENLSCTHCKERLPSHSDFLAELRKA PVVNCSTAVCORIQCDIP 1020
Db 976 RLQNTVLDWRPQVTFSENLSCTHCKERLPSHSDFLAELRKA PVVNCSTAVCORIQCDIP 1035

Qy 1021 FFGIQEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQGAFAVRSQET 1080

Db 1036 FFGIQEENATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQGAFAVRSQET 1095
Qy 1081 KVEPFEVENPLPLIVGSSVGGILLIALITAAALYKLGFFKQYKDMMSGGPPGAEPP 1137
Db 1096 KVEPFEVENPLPLIVGSSVGGILLIALITAAALYKLGFFKQYKDMMSGGPPGAEPP 1152

RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-062A-44

Query Match 58.8%; Score 3459; DB 2; Length 1163;
Best Local Similarity 60.9%; Pred. No. 9.8e-290; Indels 6; Gaps 4;
Matches 687; Conservative 142; Mismatches 294;

Qy 1 FNLDTENAMTFOENARGFGQSVVOLGSRVYVAGPOEIVAAANQSGSLYOCYSTGSCBPI 60
Db 20 FNLDTELTAPRVDGAGFGDSVVQVANGSVVVGAPOKITAAANTGGLYCCGYSTGACBPI 79

Qy 61 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQCSENTYVKGCLPFLGNSLRQPOK 120
Db 80 GLQVPPEAVNMSLGLSLAATTPSPOLLACGPTVHQCSENTYVKGCLPFLGNSLRQPOK 137

Qy 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDPRRMEKQVSTWMEOLKSKTFLSLMOYSEEF 180
Db 138 LPVSRQECPRQEDIVFLIDGSGSISSNFATMNFVRAVISQFORPSTQPSLMQFSNKF 197

Qy 181 RIHFTFEKFNQNNPRSLKIPITQLGRTHRTATGLAKVVRLEFNTINGARKNAFKILFL 240

198 QTFTEFEFRRTNPLSLASVHQLQGGFTYTATQNVVHRLPHASYGARRDATKILLIVI 257
241 TDGKFGDPLGVEDVLPEDRSGVIRYVIGVGDAPRSEKSRORLNTVASKPRDPRHVFQIN 300
258 TDGKFGDPLGVEDVLPEDRSGVIRYVIGVGDAPRSEKSRORLNTVASKPRDPRHVFQIN 317
301 NFEALKTIONLREKI FAIEGTOTGSSSSFEHMSQEGFSAATNSNGPLLSVGSYDWAG 360
318 DFDALDKIQNLQKEKI FAIEGTETTTSSSSFEHMSQEGFSAVTPDGPVLGAVGFTWSG 377
361 GVEPLTSKESKSTFNNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 420
378 GAFLYPPNNKSTFNNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 437
421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
438 QVSRQWRMKAEBVTGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 497
481 PRGQARWOCDAVL YGEGQOPWGRFGALFVLGVNGDKLTDVAIGAPGEENRGAVYLF 540
498 PRGWR-RWMCDAVL YGEGQOPWGRFGALFVLGVNGDKLTDVAIGAPGEENRGAVYLF 556
541 HGTSGSGISPSHSORISAGSLPQYFGQSGGQGLTMDGLVLDLTVGAQGHVILLRSQ 600
557 HGVLGPSISPSHSORISAGSLPQYFGQSGGQGLTMDGLVLDLTVGAQGHVILLRSQ 616
601 PVLKVKAINFEPRVARNVFECDQVYKGAZEGVRCVCHVOKSTRDLREGQOSVVT 660
617 PVLKVKAINFEPRVARNVFECDQVYKGAZEGVRCVCHVOKSTRDLREGQOSVVT 676
661 YDLALDSGPHSRAVENKNSRRTOTVGLGTOTCTETLQLOLPCIEDPVPVILRLNF 720
677 LDALDPGLSPRATQETKNSRSLRVGLGKHCENFNLLPSCVEDSVPTILRLNF 736
721 SLVGTPLSAGNLRPLVLAEDQRLFTALPFPKNCNDNICODDISITSEKSLDCLVVG 780
737 TLVGKPLAFRLNPLMAALAORYFTASLPFPKNCNDNICODDISITSEKSLDCLVVG 796
781 GRBFNVTVTVNDGDSVETQVTFPPLDLSYKYSTLQONORSNRLACESASTEV 840
797 SNLELNAEYVWVNDGDSVETQVTFPPLDLSYKYSTLQONORSNRLACESASTEV 854
841 SGALKSTSCSINHEPIFENSEVTFNTTFDVSASLGNKLLKANVTSENNMPTNKTEF 900
855 SQGTWSTSCRINELIFRGAQITFLATFDVSPKAVGLDRLLLTANVSSENNTPRTSKTF 914
901 QLELPVKYAVVYVTSKYSTKYNLFTAS-ENTSRVMOYQVSNLQBSLPSILVFLVP 959
915 QLELPVKYAVVYVTSKYSTKYNLFTAS-ENTSRVMOYQVSNLQBSLPSILVFLVP 974
960 VRLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNCVSIACVQRIQCDI 1019
975 VELNQZAVWVNDVESHQPNPSLRCSSEKXIAPPASDPLAHIQKNPVLDCSAGCLFRCDV 1034
1020 PFGIOEENATLKNLSFDWYIKTSHNLLIVSTABILLPNDSVFTLLPGQAFVRSQTE 1079
1035 PPSVQOEELDTLKNLSFCWRQILQKQVSVVSAEITFDTSVYSQLEQAFVRAQTT 1094
1080 TKVEPPEVNPPLIIVSSVGGLLALLALITAAALYKLPKPKRYQVNDME 1128
1095 TVLEKYVHNPTPLIVSSIGGLLALLALITAVLYKVGFFRQYKKNMEZ 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44
Query Match 58.8%; Score 3459; DB 5; Length 1163;
Best Local Similarity 60.9%; Pred. No. 9,8e-290;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;
QY 1 FNLDTENAMTQENARGFGQSVQVQGSVVVVGAPQEIIVAAANQSGLSYQCYSTGSCPEI 60
DB 20 FNLDTTELTAFRVDSAGFDSVVQVANSVVVVGAPQKIITAAQNGGLYQCYSTGACPEI 79
QY 61 RLOYPVEAVNMSLGLSLAATTSPPOLLACGPTVQTCENTVYVGLGCLFPGSNLRQOPK 120
DB 80 GLOVPPEAVNMSLGLSLAATTSPPOLLACGPTVQTCENTVYVGLGCLFPGSNLRQOPK 137
QY 121 FPEALRGCPQSDSOTAFILIDGSGSIIIPHDPRMEKESVTVNMEQKSKSTLFLSLMOYSEEF 180
DB 138 LPVSRQECPRQEQDIVFLIDGSGSISRNPFATMNFVRAVISQFQRPSTQFSLAQFSNK 197
QY 181 RIHFTPEFQNNPDRSLIKDITOLLGRTHATGLRKVVRELFINITNGARKNAFKILL 240
DB 198 QTHETFEFRETSPNLSLAVHQLQGGFTYTATQNVVHRLPHASYGARRDATKILLIVI 257
QY 241 TDGKFGDPLGVEDVLPEDRSGVIRYVIGVGDAPRSEKSRORLNTVASKPRDPRHVFQIN 300
DB 258 TDGKFGDPLGVEDVLPEDRSGVIRYVIGVGDAPRSEKSRORLNTVASKPRDPRHVFQIN 317
QY 301 NFEALKTIONLREKI FAIEGTOTGSSSSFEHMSQEGFSAATNSNGPLLSVGSYDWAG 360
DB 318 DFDALDKIQNLQKEKI FAIEGTETTTSSSSFEHMSQEGFSAVTPDGPVLGAVGFTWSG 377
QY 361 GVEPLTSKESKSTFNNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 420
DB 378 GAFLYPPNNKSTFNNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPRYOHIGLVAMFR 437
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
DB 438 QVSRQWRMKAEBVTGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 497
QY 481 PRGQARWOCDAVL YGEGQOPWGRFGALFVLGVNGDKLTDVAIGAPGEENRGAVYLF 540
DB 498 PRGWR-RWMCDAVL YGEGQOPWGRFGALFVLGVNGDKLTDVAIGAPGEENRGAVYLF 556

Qy	541	HGTSGSGISPSHORIAGSKLSPRLQYFGQSLSGGDLTMDGLVLDLVGAQGHVLLRSQ	600
Db	557	HGVLGPSISPSHQRIAGSDLSRLQYFGQALSGGDLTQDGLVLDLVAGARGQVLLRTR	616
Qy	601	PVLRFVKAIMEFNPREVARNYFECNDQVVKCKEAGEVRVCLHVOKSTDRRLREGQIQSVVT	660
Db	617	PVLWVGVMQFIPAEIIPRSFAFECREQVWSEQTIVQSNICLYIDKRSKNLIGSRDLQSSVT	676
Qy	661	YDLALDSGRPHSRAVNETKNSSTRQTVLGLTCTCETLKLQLPNCIEDPVSPIVLRNF	720
Db	677	LDLALDGRUSPRATFOETKNRSLRSVRVLGKAHCENFNLLEPSCVEDSVITTLRLNF	736
Qy	721	SLVGTPLSAFNGRLPVLAEADAORLFTALFFPKKCGNDNICODDLSITFSFMSLDCLVVG	780
Db	737	TLVGKPLLAFLNLRPLAALAAQRYFTASLPFEKNCAGADHICQDNLGISFSFPGKSLVVG	796
Qy	781	GPREFNVTVVRNDGDSYFTQVTFPPPLDLSYRKYSTLQONORSQSRWLACESASSTEV	840
Db	797	SNLELNAEVVRNDGDSYGTITTFSHPAGLSYRYAEGQKQQLRSLHLTCDSPVG--	854
Qy	841	SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKXNVTSNNMPTNKTEF	900
Db	855	SQGTWSTSCRINHLIFRGGAQITFLATFDVSPRAVLGDRLLLTANVSSNNTPRTSKTTF	914
Qy	901	QLELPVKYAYVMVTSHGVS TKYLNFTAS-ENTSRVWMOHOYOVSNLQORSIPISLVFLVP	959
Db	915	QLELPVKYAYVTWSSHEQTKYLNFSBSEESHVAMHRYQVNNLQORDLPVSNFWVP	974
Qy	960	VRNLOTVIDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDI	1019
Db	975	VELNQEAVMVDEVSHFQNTSLRCSSEKIAPPASDFLAHQKPNFLDCCSIAGCLRFCDV	1034
Qy	1020	PFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVFTLLPQGAFAVESQTE	1079
Db	1035	PSFSVQELDFTLKENLSFGWVRQILQKQVSVSVABITFTSVYSQLPQGEAFMEAQTT	1094
Qy	1080	TKVEPPEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKROYKDMNSE	1128
Db	1095	TVLEKYKVHAPTPLIVGSSIGLILLALITAVLYKVGFFFRQYKEMMEZ	1143

Search completed: June 7, 2004, 17:19:23
 Job time : 22.3484 secs

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds
(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAWTQENARGFGQ.....PKRQYKMMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	1 RWHU1B	cell surface glyco
2	4470	76.1	1153	2 S00551	leukocyte surface
3	3479	59.2	1163	1 RWHU1C	cell surface glyco
4	1553.5	26.4	1170	2 S03308	cell surface glyco
5	1538.5	26.2	1163	2 S03126	lymphocyte fuction
6	1153	19.6	1179	2 A53213	integrin alpha-8 c
7	1103.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1095	18.5	1170	2 I45914	integrin alpha-2 s
9	1071	18.2	1178	2 S44142	VLA-2 protein homo
10	1068	18.2	1181	2 A33998	integrin alpha-2 c
11	1062	18.1	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	635	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha-4 c
16	579.5	9.9	1054	2 JC7294	alphan integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	542.5	9.2	1034	2 A31008	integrin alpha-v c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	531.5	9.0	1072	2 A38457	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	8.9	1048	2 A27421	integrin alpha-6 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

30 496 8.4 1394 2 A29637 position-specific
31 495.5 8.4 1039 2 A34269 integrin alpha-2b
32 494 8.4 1137 2 JCS950 integrin alpha-7 c
33 490.5 8.3 1037 2 A60163 glycoprotein IIB -
34 488 8.3 1135 2 I61186 alpha-7 integrin -
35 486 8.3 126 2 B30892 leukocyte adhesion
36 480 8.0 1106 2 S38783 integrin alpha cha
37 467 7.9 1226 2 S44824 p54f2.1 protein -
38 453 7.7 1045 2 S60571 integrin alpha v c
39 445.5 7.6 1139 2 S28277 hypothetical prote
40 423 7.2 1115 2 T09403 integrin alpha cha
41 418.5 7.1 1115 2 T09433 integrin alpha cha
42 391 6.7 764 2 T36916 glycoprotein IIB -
43 309.5 5.3 1086 2 T18523 integrin alpha cha
44 299 5.1 604 2 T36917 glycoprotein IIB -
45 298.5 5.1 272 2 A55348 integrin alpha-1 -

ALIGNMENTS

RESULT 1
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M
eukocyte integrin alpha chain; neutrophil adherence receptor alphas chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text change 31-Dec-2000
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I58567
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C
B.
A:Reference number: A31108; MUID:88315033; PMID:2457584
A:Accession: A31108
A:Molecule type: mRNA
A:Residues: 1-1153 <OR>
A:Cross-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148
A:Note: part of this sequence was confirmed by protein sequencing
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor 1
A:Reference number: A28915; MUID:88257215; PMID:2454931
A:Accession: A28915
A:Molecule type: mRNA
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA59
A:Note: the authors translated the codon TAC for residue 1129 as Thr
A:Note: part of this sequence, including the amino end of the mature protein, was confi
R:Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A:Reference number: A41600; MUID:92073318; PMID:1683702
A:Accession: A41600
A:Molecule type: DNA
A:Residues: 1-9 <SHE>
A:Cross-references: GB:W76724; NID:G180018; PIDN:AAA58410.1; PID:G553215
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes
A:Reference number: A94193; MUID:88190151; PMID:2833753
A:Accession: A30892
A:Molecule type: mRNA
A:Residues: 917-1042 <AR2>
A:Cross-references: GB:M18044
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A:Reference number: A32218; MUID:89098893; PMID:2563162
A:Accession: A32218
A:Molecule type: mRNA
A:Residues: 9-1153 <HIC>
A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A>Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.P.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; PMID:93123748; PMID:8419480
A:Accession: A46526
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FLB>
A:Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
R;Pierce, M.W.; Remold-O'Donnell, B.; Todd III, R.P.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PIE>
A:Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A>Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:G180184; PIDN:AA51960.1; PID:G553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	76
QY	61	RLOPVEAVNNSLGLAATSPOLLACGTPVHOTCSNTYVKGCLFLPGSNLRQPOK	120
DB	77	RLOPVEAVNNSLGLAATSPOLLACGTPVHOTCSNTYVKGCLFLPGSNLRQPOK	136
QY	121	FPFALRGCPQSDSIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTLFSLMQYSEEF	180
DB	137	FPFALRGCPQSDSIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTLFSLMQYSEEF	196
QY	181	RIHFTFKFQNNPNSRLIKITQLLGRTHATGIRKVVRELFNITGARKNAEKILILI	240
DB	197	RIHFTFKFQNNPNSRLIKITQLLGRTHATGIRKVVRELFNITGARKNAEKILIVI	256
QY	241	TGDEKFGDPLGVEDVIPADREGVIRYVIGVDAPRSEKSRQELNTVASKEPRDHVQIN	300
DB	257	TGDEKFGDPLGVEDVIPADREGVIRYVIGVDAPRSEKSRQELNTIASKEPRDHVQVN	316

QY	301	NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVGSYDWAG	360
DB	317	NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVILGARYQHIGLVAMER	420
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVILGARYQHIGLVAMER	436
QY	421	QNTGHWESNANVKGTCIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETEGRGQSVSCPL	480
DB	437	QNTGHWESNANVKGTCIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETEGRGQSVSCPL	496
QY	481	PRGQARWQCDVAVLYGEGQPMGRFGAALTIVLGDVNGDKLTVAIGAPHEENRGAIVILF	540
DB	497	PRGQARWQCDVAVLYGEGQPMGRFGAALTIVLGDVNGDKLTVAIGAPHEENRGAIVILF	556
QY	541	HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSSGGQDLTMDGLVLTVAQGHVLLLRSQ	600
DB	557	HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSSGGQDLTMDGLVLTVAQGHVLLLRSQ	616
QY	601	PVLRVKALMEFNPFEVARNVFECDQVVKKGAGEVRVCLHVQKSTRDLRLRGGQIOSVVT	660
DB	617	PVLRVKALMEFNPFEVARNVFECDQVVKKGAGEVRVCLHVQKSTRDLRLRGGQIOSVVT	676
QY	661	YDLALDSGPHSRVFNENKSTRQTVLGTTCETIKLQLPNCIEDPVSPVILRLNF	720
DB	677	YDLALDSGPHSRVFNENKSTRQTVLGTTCETIKLQLPNCIEDPVSPVILRLNF	736
QY	721	SLVGTPLSAFGNLRPVLAEDAQLFTALPFFPKXNCGNDNICQDDLSTFSPMSLDCLVWG	780
DB	737	SLVGTPLSAFGNLRPVLAEDAQLFTALPFFPKXNCGNDNICQDDLSTFSPMSLDCLVWG	796
QY	781	GRPEFNVTVTRNDGSDSYRTQVTPPEPLDLSVRKVTSTONORSORSLACESSASTEV	840
DB	797	GRPEFNVTVTRNDGSDSYRTQVTPPEPLDLSVRKVTSTONORSORSLACESSASTEV	856
QY	841	SGALKSTSCSINHPIIPENSEVTFNITDQVDSKASLGKLLKLLKANVTSENNMPTNKTEF	900
DB	857	SGALKSTSCSINHPIIPENSEVTFNITDQVDSKASLGKLLKLLKANVTSENNMPTNKTEF	916
QY	901	QLELPKYAVMVVTSHGVSSTKYLNTASNTSRVMOHOYQVSNLQORSIPISLAVLPV	960
DB	917	QLELPKYAVMVVTSHGVSSTKYLNTASNTSRVMOHOYQVSNLQORSIPISLAVLPV	976
QY	961	RLMOTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCIAVCORIQDIP	1020
DB	977	RLMOTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVAVNCIAVCORIQDIP	1036
QY	1021	FFGIQBEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDVSFTLLPGQCAFVRSQTET	1080
DB	1037	FFGIQBEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDVSFTLLPGQCAFVRSQTET	1096
QY	1081	KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMHSEGGPPGAEQ	1137
DB	1097	KVEPFEPVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMHSEGGPPGAEQ	1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C;Accession: S00551; 159078
R;Pytela, R.
EMBO J. 7, 1371-1378, 1988
A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the ;
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PPT>
A;Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983
A;Note: the authors translated the codon CAC for residue 569 as Gin

[illegible]

Query Match	59.28;	Score 3479;	DB 1;	Length 1163;
Best Local Similarity	61.14;	Pred. No. 1.3e-233;		
Matches	690;	Conservative 141;	Mismatches 292;	Indels 6;
				Gaps 4;
Qy	1	FNLDTENAMTFQ--ENARGFGSVVQLQGSRRVVVGAPOEIVAAVORGSILYQCDYSTGSCPEI	60	
Db	20	FNLDTTELTAFRVDSAGFGSDSVQYANVWVGAPOKITTAANQTGGLYQCGYSTGACPEI	79	
Qy	61	RLQVPVAVNMSGLSLAATTSPPOLLACGFTVHOTCSENTYVKGCLFLGSLNLRQPOK	120	
Db	80	GLQVPPVAVNMSGLSLAATTSPPOLLACGFTVHOTCSENTYVKGCLFLGSLNLRQPOK	137	
Qy	121	FPEALRCPOBDSIAFLVDGSGSIIIPHDFFRAKEPFTSTVMEOLKSKTILFSLMOYSEEF	180	
Db	138	LPVSRQSCPBQODIVFLIDGSGSISSENATMWNFVRAVISQFORSTQFSLMOFSNKP	197	
Qy	181	RIHTTFEPONNPRSLRIPITOLGORTATGIRKVVRELTNITNGAKNFAKILILI	240	
Db	198	QTHETFEPRTSNPLSASVHQLQFTTATAIONVHRLPHASGARDRATKLIVI	257	
Qy	241	TGCKFGDELGYEIVPEADREGVIRVVGDAFRSEKSKOELNTVASPPRDRHVQIN	300	
Db	258	TGCKKSGSLDYKDVPMADAAGIIRVAIGVLAFORENSWKENLDIASKPSQEHFKVE	317	
Qy	301	NFEALKTIONLRKIPAIKGTQGTSSSEHEHMSQSGPAAITSNGPLISTVGSYDWAQ	360	
Db	318	DFDALKDIONLKEIPAEIETETTSSTSSPELEMAQGFSAVPTPDGVPVLCGAVGSFTWSG	377	
Qy	361	GVPLVTSKESKSTFNMTREVSMDKDAVLGAAAIILNRVQSLVIGAPRYCHIGLVAMER	420	
Db	378	GAFILPPNMSPTFLNMQENVMDRDSVLGISTELALWKGVQSLVIGAPRYCHIGLVAMER	437	
Qy	421	QNTQMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSCPL	480	
Db	438	QVSRQWRKAEVCTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSCPL	497	
Qy	481	PRGORAWCDAVLVGOGOPWGFAGALTVLGVGVGDKLTDVAIGAPGEDNRGAYLFP	540	
Db	498	PRGWR-RWMDAVLYGQGHPPWGFAGALTVLGVGVGDKLTDVWVIGAPGEENRGAYLFP	556	
Qy	541	HGTSGSGTSPSHSQRISAGSKLSPRLQYFGOSLGGQDLTMDGLVDLVVGAQGVHLLRSQ	600	
Db	557	HGVLGPSISPSHSQRISAGSKLSRLQYFGQALSQGLTQDGLVDLVVGAQGVHLLRTR	616	
Qy	601	PVLVKALMEPNPREVARNPECHDQVVKGEAGEVRLVHVKQSTDRLEBQIQSVIT	660	
Db	617	PVLVGVSMQPIPAEIPRSAPCEBQVSEBTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676	
Qy	661	YDLALDGRPHRAVFNKTSNRQTVGLTQTCETLKLQPNCHIEDPVSPITVLRLNF	720	
Db	677	LDLALDGRSLSPRAFTQETKNSLRVRLGLKAHCENFNLLPSCVEDSVPTITLRLNF	736	
Qy	721	SLVGTPLSAFNLAPVLABDQRLFTALPPFERKVCNDNITCQDDLSITFPMSLDCLVWG	780	
Db	737	TLVGKPLLAFLNRLPMLAADAQRYFTASLPFERKNCGADHICQDNLGSPFGLKSLVG	796	
Qy	781	QPRSFNTVTVRNDEGDSYRQVTFPPPLDLSYRKVSTFLQNRQSRWRLACRSASSTEV	840	
Db	797	SNLEHAEVWVNDGDSYGTITFTSPAGLSYRYVAGQYQGLRLSHITCDSAPVG--	854	
Qy	841	SGALKSTSCSINHPIPPENSEVFTNITFDVDSKASLGNKLLKANVTSENMPRTNTEP	900	
Db	855	SGQWSTSCRINHILFRGGAQITFLATFDVSPKAVLGDRLILTANVSSENTPRTSKTF	914	
Qy	901	QLELPVKYVAVNTSHCVSTKYLNFAS-ENTSRVMCHQYOVSNLQORSILPSLVLVP	959	
Db	915	QLELPVKYAVTVVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQORDLPVSINFWVP	974	
Qy	960	VRLNQTIWDRPQVTFSENLSSTCHTERLFPSSHDFLAELKAPVWNCISIAVCOICDI	1019	
Db	975	VELNQEAVWVDSVHPQNSLRCSEKIIAPPASDFLAHQKPNVLDCSIAGCLRPRCDV	1034	
Qy	1020	PPFGIOBEFNATLKNLSFDVYIYKSHNHLIVSTABILFNDVSTFILLPQGAQVRSQTE	1079	

Db	1035	PSFSVQVEELDFTLKNLSFGWVRQLQKXSVSVVAEITFTDTSVYSQLPQGEAFWRAQTT	1094	
Qy	1080	TKVEPEFVNPPLPLIVGSGVGLLILALITAAALYKLGKFFPKRYKQMMSE	1128	
Db	1095	TVLEKYKVNPTPLIVGSGSISGGLLILALITAVLYKVGFPKRYKQMMSE	1143	
RESULT 4				
S03308				
cell surface glycoprotein CD11a precursor - human				
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function				
C;Species: Homo sapiens (man)				
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 20-Aug-1999				
C;Accession: S03308; A47565; A47565; A48759; S36044				
R;Larson, R.S.; Corbi, A.L.; Beruan, L.; Springer, T.				
J. Cell Biol. 108, 703-712, 1989				
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit				
A;Reference number: S03308; MUID:89139587; PMID:2537322				
A;Accession: S03308				
A;Molecule type: mRNA				
A;Residues: 1-1170 <LFA>				
A;Cross-references: EMBL:X00796; NID:931421; PIDN:CAA68747.1; PID:931422				
A;Note: part of this sequence was confirmed by protein sequencing				
R;Cornwell, R.D.; Gollan, K.A.; Hickstein, D.D.				
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993				
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr				
A;Reference number: A47458; MUID:93248261; PMID:8097887				
A;Accession: A47458				
A;Molecule type: DNA				
A;Residues: 1-20 <COR>				
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)				
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.				
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993				
A;Title: Identification of cell-specific and developmentally regulated nuclear factors				
A;Reference number: A47565; MUID:93281759; PMID:8099450				
A;Accession: A47565				
A;Molecule type: DNA				
A;Residues: 1-20 <SHE>				
A;Cross-references: GB:M95609				
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.				
J. Biol. Chem. 268, 19305-19311, 1993				
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.				
A;Reference number: A48759; MUID:93374910; PMID:8103515				
A;Accession: A48759				
A;Molecule type: DNA				
A;Residues: 1-20 <NUE>				
A;Cross-references: EMBL:222804; NID:9311405; PIDN:CAA80461.1; PID:9311406				
C;Genetics:				
A;Gene: GDB:ITGAL; CD11A				
A;Cross-references: GDB:119757; OMIM:153370				
A;Map position: 16p11.2-16p11.2				
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom				
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr				
F;1-25/Domain: signal sequence #status predicted <Sig>				
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted				
F;154-317/Domain: von Willebrand factor type A repeat homology <VMA2>				
Query Match 26.4%; Score 1553.5; DB 2; Length 1170;				
Best Local Similarity 34.7%; Pred. No. 1.7e-99;				
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;				
Qy	1	FNLDTENAMTFQ--ENARGFGSVVQLQGSRRVVVGAPOEIVAAVORGSILYQCDYSTGSCB	58	
Db	26	YNLDVRGASFSPPRAGRHFGRVLQV-GNGVTVGAPGE--GNSTGSLYQCSGTGCHL	81	
Qy	59	FIELOVPVAVNMSGLSLAATTSPPOLLACGFTVHOTCSENTYVKGCLFLGSLNLR--	115	
Db	82	PVTLR-GSNYSKYLQWTLATDFTGSLIACDPGLSRTCDQNTYISGLCYLFRNLQGP	140	
Qy	116	-QOQKFPBALRGCPQEDSIAFLVDGSGSIIIPHDFFRAKEPFTSTVMEOLKSKTILFSLM	174	
Db	141	LQGRPGQBCING---NVDLVFLDGSMSLQDEPKILDFMKDVMKKLSNITSYQFAAV	196	

175 OYSEEFRIHFTKEPQNNPNSRLKPIQTQALLGRTHATGIRKVVRELFNITNGARKNAP 234
 197 QSTSTKTFDSYVYKWDQDALLKXVHMLLNTFGAINVATEVREBELGARPDAI 256
 235 KILILITDGEKFDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPED 294
 257 KVLIIITIDGE--ATDSGNIDAAD-----IIRVIGIGKHQFQKESQETLHKFASKPASE 309
 295 HYPQINNPEALKTIQNLQEKIPIAIEGTOTGSSSEFHEMSQEGFSAALITNGPLLTSTVG 354
 310 FVKILDTFBLKDLFTLQKKIYVIEGTQKDLTSPNMLSSGSIADLSRSHGVAVGAG 369
 355 SYDWAAGGVF-LYTSKEKSTFINNTRVDSNDVAYLGYAAA-IILNRVQSLVLGAPRYOH 412
 370 AXDWAGGFLDKADLQDDTFIGNPELTPVRAGYLGTYVWLPSPKTKTSLASGAPRYOH 429
 413 IGLVAMFR--QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOT 470
 430 MGRVLLFQBPQGGHWSQVQTHGTQIGSYFGELCGVDVDQDGETELLIGAPLYGSG 489
 471 RGGQSVCPPLPRQARWOCDAV--LYGQGGQPGWGRFGAALTVLGVNGDKLTDVAIGAP 528
 490 RGRGVFIY-----QRQLGFEEVSELOQDPGYPLGRFGEAITALTIDNGDLVDVAVGAP 544
 529 GEDNRGAVYLPHGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDQLTMDGLVDLTV 588
 545 LEE--QCAVYVFNGRHG-GLSPQSPQIRGTQVLSGIOWFGRSINHGVKDLGDLADVAV 601
 589 GAQGHVLLRSQPLVAKVKAIMEPNPREVARNVPCNDQVV-KKKGAGEVRVCLHVKQSPR 647
 602 GAGSQMIVLSRPVDMVIMLSPSPAEIPVHEVEGYSYSTSNKRGVNTICFQI-KSLY 660
 648 DLREBQIOQSVVYTDIALDSGRPHSAVFNENKNSRRQTVGLTQTCTELKQLPNCI 707
 661 PQP-QGRVANLTYTTEQLDGHRTERRGLFPFGGRHLLRNIAVT-TSMSCDPSFHPVGV 718
 708 EDPVSPVILRLNESL---VGTPLS--AFGN-----LRPVLAEDAQLFTALPFFKNGCN 757
 719 QDLISPINVSLNPSLMEERGGTPRDQQAQKQIPILRPSLHSETWBI-----PFEKNCEB 773
 758 DNIQDQDLSTPMSLDCVLVGGPREFNVTVTRNDGDSYRTQVTPFPFLDLSRYKVS 817
 774 DKKEANLVSPSPASRALRLTAPASLVELSLNLEADYVWQLDLHFPPLSFRKVE 833
 818 TLQNRQSRWLACHS--ASSTEVSGALKSTSCSINHIPPENSEVTNITFDVDSKAS 875
 834 ML---XPHSQIIVSCBELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSS 887
 876 LGNKLLKANVTSENN---MPTNKTETQLLPVKYAVVWVTSHGVSSTKYLNTFASGN 931
 888 WEDSVELHANVTCNEDSDLEDNSATTI---IPILYPINILIQCEDSTLYVSTPKGP 944
 932 TSVVMQHQYOV---SNLQORSIP-LSLFLVPVRLNQTVINDRPQVTFSENLSSTCHK- 986
 945 KIQVKNMTQVRIQPSIHQNTIPTLEAVVGVQPPSEGPITHQWSVQMEPPV--PCHVED 1002
 987 -RLPSHSD--FLAELRKAPVNCISAVCQRQCQDIPFGIQEENFATLKNLSFDWIK 1043
 1003 LERLPDAAEPCPLGALFRCPVW-----PFQELVQVIGLLELVGEIE 1044
 1044 TSHNHLIVSTABILFNDSTVFTLLPQGGAFVRSQETKVEPPEVENPFLPIVGGSVGGL 1103
 1045 AS-SMPSLCSLSISNSKHFHYLGSNASL-AQVVMKVVDVYVYKQMLYLYVLSIGIGLL 1102
 1104 LLALITAAIYKLGFPKQYKDMSEG-GPPGAP 1136
 1103 LLLLLIFVLYKVGFFKRNLEKXMEAGRGPVNGIP 1136

RESULT 5
 I56126
 lymphocyte fuction-associated molecule-1-alpha - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
 C:Accession: I56126
 R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
 J. Immunol. 147, 369-374, 1991
 A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
 A:Reference number: I56126; MUID:91268576; PMID:2051027
 A:Accession: I56126
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1163 <RBS>
 A:Cross-references: GB:M60778; MID:g198785; PIDN:AAA39426.1; PID:g198786
 C:Gene: LPA-1
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 F:151-315/Domain: von Willebrand factor type A repeat homology <VMA>

Query Match 26.2%; Score 1538.5; DB 2; Length 1163;
 Best Local Similarity 34.2%; Pred. No. 1.9e-98;
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

QY 1 FNLDTENANTFOENA-RGFQSVVQLQSGSRVVGGAPOBIAVANDRGSLVQCQDYSTGSCPEP 59
 DB 24 YNLDTPTQSFQAQGRHFGYQVLIQEDG-VVVGARGE--GDNTGGLYHCTSTSEFCQP 79
 QY 60 IRLQVPEAVNMSLGLSLAATTSPPQLACQPTVHQTCSNTYVYKGLCFGLFSNLRQOPQ 119
 DB 80 VSLH-GSNHTSKYLMTLATDAKGSLLACDPLGSLRTCDQNTYLSGLCYLFPQSLGPM 138
 QY 120 KPPALRCQPCQDSDIAFLVDGSGSIIIPDHERRAKERISTVMEOLKSKTKTFLPSLMQYSE 179
 DB 139 QNRPAYQCMKGVLDVFLFDGQSLDRKDEKILEFMQVWRKLSNTSYQFAAVQFSD 198
 QY 180 FRIHFTKEF-QNNFNPRLSLIKPIQTQALLGRTHATGIRKVVRELFNITNGARKNAFKILI 238
 DB 199 CRTEPTFLDYVQKNKPNPDVLLGSVQPMFLLENTFRAINVVAHVFKESGARPDAIKVLV 258
 QY 239 LITDGEKFDPLGYEDVPEADREG-----VIRYVIGVGDAPFRSEKSRQELNTVASKP 291
 DB 259 IITDG-----EASDKGNISAADHITRYIIGIKGFVSVQKQTLHIFASEP 304
 QY 292 PRDHVQINNPEALKTIQNLREKIFAIEGTQTSSTSEFHEMSQEGFSAALITNGPLLS 351
 DB 305 VEEFVKILDTLQEKDLPTDLQRIYAIIEGNNRQDLTSFNMLSSGSIADLSKGVAVG 364
 QY 352 TVGSYDWAAGGVF-LYTSKEKSTFINNTRVDSNDVAYLGYAAA-IILNRVQSLVLGAPR 409
 DB 365 AVGAKDWAGGFLDLREDLQCATFVQCEPLTSQVREGGYLGYTVAWMTSRSSRPLAAGAP 424
 QY 410 YOHIGLVAMFR--QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 467
 DB 425 YQHVQVLLFOAPAGGEWNTQKIEGTQIGSYFGELCSVDLQDQGEAEILLIGAPLP 484
 QY 488 BQTRGGQSVCPPLPRQARWOCDAVLYGEGQPGWGRFGAALTVLGVNGDKLTDVAIGA 527
 DB 485 GEQRGRVFTY---QRRQSLPEAVVSELQGDPGYPLGRFGEAITALTIDNGDLTDVAIGA 541
 QY 528 PGEEDNRGAVYLPHGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDQLTMDGLVDLT 587
 DB 542 PLEB--QCAVYVFNGRHG-GLSPQSPQIRGTQVLSGIOWFGRSINHGVKDLGDLADV 598
 QY 588 VCAQGHVLLRSQPLVAKVKAIMEPNPREVARNVPCNDQVVKGEAG-EVRVCLHVKQST 646
 DB 599 VCAEGRVVVLSRRPVDVWVTELSFSPBEIPVHEVEGYSYSAEEQKHGKVLKACPRFKPL 658
 QY 647 RDLREBQIOQSVVYTDIALDSGRPHSAVFNENKNSRRQTVGLTQTCTELKQLPNC 706
 DB 659 PQ--FOGRLANLSTYQLDGHMRMRSLGFLPDGGSHLSGNTSITTP-DKSLDFFHFPD 715
 QY 707 IEDPVSPVILRLNESL---VGTPLSAGFN-LRPVLAEDAQLFTALPFFKNGCNDNICQ 762
 DB 716 IQDLISPINVSLNPSLMEERGGTPRDQQAQKQIPILRPSIHTV-TKEIPFEKNCGEDKKE 774

```

QY 763 DLSITTFMSLDCLVGGP-----RBNVTVTVVRNDGEDSVRTQVTFPPFLDLVSRKV 816
DB 775 ANLTSSPARS-----GFLRLMSSASLAVETILNSGEDAYVWVRLDLDLDFPRGLSFRKY 827
QY 817 STLQNRQSRWELACESASTEVSGAL-KSTSCSINHPIFPENSESVTENIPFVDVSKAS 875
DB 828 EMLO-----PHSRMPVSCSEL--TEGSELLTKYLCNVSSPIFKAGQEVSLQVNFILNSS 882
QY 876 LGKLLKLVKAVTSEN-NMPTNKTEPQLBELPVKYAVYVWVTSHGVSSTKLTANTASNTSR 934
DB 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNVILTKQENSTLYISFTPKGPKTKQ 942
QY 935 VVOHOYQVNSLGNORSIPISLVLVPLVRLNQTVWDRPQ-----VTESENLS-----TCHTK 986
DB 943 QVOHVTVQ-----RIQPSAYDENMPT-LEALVGVPRPHSEDLITITWSTQDPLVTCHE 996
QY 987 E-RLPSEHDFLAELRKAPVNVCSIAVCQRIQCDIPFPGIQEENFATLKNLSFDMWYIKTS 1045
DB 997 DLKRPSE--AEQPCLPV-----QFRCPIVE--RWEILIQVTVGVLSKEIKAS 1042
QY 1046 HNHLLIVSTAEILFNDSVETILLPCQAFVRSQTETKVPPEVNPVPLPLVGVSSVGLLILL 1105
DB 1043 -STLSLCSLSVSFNSKHFHLYGSKA-SEAQVVLVVDLIHKEMLHVTVLSGIGELVILL 1100
QY 1106 ALITAAALYKLGFPKROVKDM-SEGPPGGAEP 1136
DB 1101 FLIFALYKVGFPKRLKEKMRADGCVNGSP 1132

RESULT 6
A33213
Integrin alpha-E chain - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C/Accession: A53213
R/Shaw, S.K.; Cepce, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit.
A:Reference number: A53213; MUID:94164962; PMID:8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: GB:L25851; NID:9457244; PID:9457245
C:Genetics:
A:Gene: ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
A:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1153; DB 2; Length 1179;
Best Local Similarity 29.1%; Pred. No. 1.3e-71;
Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;

QY 45 GSIYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPQLLACGPTVHQ 95
DB 65 GPLHRCSLVODEILCHPVEHVPVPKGRHGVTVVRSHEHGLVLCI-----QVLVRRP--HS 117
QY 96 TCSSENTVVKGLCFGLFSNLRQQPQ-----GTCSLGPDLEPQAQAFDLENLLDPAQVDTGDCVSNKEGGEDDVNTA 174
DB 118 LSELT---GTCSLGPDLEPQAQAFDLENLLDPAQVDTGDCVSNKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAPLVDSGSIIPHDPRFAKEPISTVMEQL-- 164
DB 175 RQRALKEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEED 234
QY 165 KSKTILPSLMQYSEERFIHFTTFKFNQNNPNRSLIKPIQTQLGRTHTATGIRKRVVRELFN 224
DB 235 KCECNFALVQYGVQVITQTEFDRSDQVMAASLARVQNIQVGSVTKTASAMQHVLDSITF 294
QY 225 ITNGARNAFKIILITDGEKFGDPLCYEDVIEADREGVIRVIVGVGDAFSEKSRQL 284

```

```

DB 295 SSHGSRKASKVMVVLTDGGIFEDFLNLTTVINSPKQGVVERPAIGVGEERFKSARTAREL 354
QY 285 NTVASKPPDRHVQNNPEALKTIONQLREKIFAIEGTQTGSSSSPEHEMQEGEFAAIT 344
DB 355 NLIASDPDETAFKVTWYVMDLGLLSKLYVILISMEGT---VGDALHYQLAQIGFSAQIL 411
QY 345 SNGP-LLSTVGSYDWAGGVFLY--TSKEKSTPINNTRVDSMDMDA---YLGAAAIILRN 398
DB 412 DERQVLGAUGFRFWSGGALLYDTRSRGRFLNQTAARAADEAAQAYSLGYAVAVLHKT 471
QY 399 RVQSLVLGAPRYQHIGLVAMFR-QNTGWMESNAV-KGTQIGAYFGASLCSVDVDSNGST 456
DB 472 CSLSVAGAPQVKKHG--AVPELQKGRASFLPVEGEQMGSYFGSELCLPVDIMDGST 529
QY 457 DLVLIGAPHYEQYFGQVSVCPPLPRQARWQCDVLYGQGPWGFRGGAALTVLGVDN 516
DB 530 DFLVAAFPYHVGEGEVYVRLSB-QDGSFSLARILSHGPGTNRARFGFAMAGDLS 588
QY 517 GDKLTDAIGAIP-----GEEDNR--GAVLPHGTSGSISPSHSQRIAGSKLSPLRQYFGQ 570
DB 599 QDKLTDAIGAIPLEGFAGDQASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
QY 571 SLSGQDUTMDGLDVLTVGAQGHVLLRSQVLAVKALMEFNPREVARNVFECNDQVYKG 630
DB 648 SMAGGFDISGDLADITVTLGAQVVRPRSRPVRLKVSMAFTPSALP-----LGP 697
QY 631 KEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRVAFNPKNSTRQTQVL 690
DB 698 NGVNVRLCFEI-SSVTTASESGLREALNFTLDVQKQRRRLQCSDVRSCLGCLREWS 756
QY 691 GLTQTCETLKLQLEN-----CHIEDPVSPVLRLNLSVGTPLSAFGNLRPLVAEDAQRFLT 746
DB 757 SGSQLCEDL-LMPTEGELCEEDCFNSASRVSYQL-QTPEGQTDHPQPIIDRYTEPAI 814
QY 747 ALFPFKKNCNDNICODDLSITTFMSLDCLVGGPREFNVTVTVRNDGEDSVRTQVTF 806
DB 815 FQLPYKAKCKKLFCVLAELQLA-TTVSQOELVGLTKELTLNINLTNSGEDSVYMTSMALN 873
QY 807 PPLDLSYKAVSTLQNRQSRWRACBSASTEVSGALKSTCSINHPIFPENSESVTENI 866
DB 874 YPRNLQ-----LKMOKPPSPNIQCDPQPV--ASVLMNCRIGHPVL-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKULLKANVTSENN---MPTNKTEPQ---LELFPYAVVWVTSHG 919
DB 924 VQULEENAPNRTADITVTVTNSNERSLANETHLQFRHGFVAVLSKPSIMYVNTGQL 983
QY 920 S--TKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVLVPLVRLNQTVWDRPQVTFSE 977
DB 984 SHHKEFLPHVHGEN--LFGAEYQ-----LQICVPTKLRGLQVAAVKCLTRTQ 1028
QY 978 NLSSTCHTKERLPSEHDFLAELRKAPVNVCSIAVCQRIQCDIPFPGIQEENFATLKNLS 1037
DB 1029 ASTVCTWSQERACAYSS-VQHVEMHVSVCVIA-----SDKENVTVAABIS 1073
QY 1038 PDWYIKTSNHLILVST-----AEILFNDSVFTLLPQGGAFVRSQTKETKVEPPEVNP 1091
DB 1074 WD-----HSELLKDVTELQILGEISFNKSLYEGINAENH-RTKITVWFLKDEKYHSL 1125
QY 1092 PLTVGSSVGGILLALITAAALYKLGFPKROVKDMSE 1128
DB 1126 PIINGSVGGLLVILVILVILFKGFPKRYQOOLNLE 1162

RESULT 7.
Integrin alpha-1 chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
C/Accession: A45226
R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit
A:Reference number: A45226; MUID:9315124; PMID:8428973

```

A:Accession: A45226
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1151

 A:Experimental source: hepatoblastoma cell line HepG2
 A>Note: sequence extracted from NCBI backbone (NCBIP:124326)
 F:142-317/Domain: von Willebrand factor type A repeat homology <WAL>

Query Match 18.8%; Score 1103.5; DB 2; Length 1151;
 Best Local Similarity 27.8%; Pred. No. 3.6e-68;
 Matches 344; Conservative 212; Mismatches 486; Indels 195; Gaps 44;

```

QY 1 FNLDTENAMTQ-ENARG-FGQSVQL---QGSRVVVGAPQEIIVANQSGSLVQDYSGS 56
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 FNVVDKNSMTFGSPVEDMFGYTVQYENEEGKWLLGSLVQGPKNRTGDTVKCPVGGE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 CSP-IRLOVPVEA-----VNMSLGLSLAATSPPOLLAGCPVTHQTCSENYVRGL 106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 SLPCVKLDELVNTSIPNTVEKNTFTGSL-VTPNMGFLAGCPLYAYRGHLHYTTGI 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 CFLFGSNLRQOQKPFPEARLGCQSDSDIAFLVDCSGSIIIPHDFFRAKEFTSTWELQK- 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 120 CSDVSPPTFOVNSIAP--VOECSTQ-LDITVIVLDCGNSIYPWDSVTA--FLNDLLKMDI 174
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 -KSKTLFSLMOYSEBEFIHFTKEFQNNPNRSLIKPIQILGR-THTATGIRKVVREL 223
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 175 GPKQTOGVIVQGENWTFHFNLNKYSSTEEVLVAAKIVQRCGRQTMALGTDTRKEAP 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 224 NITNGARKNAKPLILITDGEKFGPLGYEDVIPEDRGVIRYVIGVGDAPF-----SE 278
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 235 TEARGARRGVKVMVITVDGESH-DNHRLLKVIQCDENIQRFSAIILGSTRNGLSTE 293
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 279 KSRQELNTVASKPRPDHVFQINNPEALKTIONLREKIPAEIGTGTGSSSEHEMSQEG 338
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 294 KVEETKSTASRTEKPHFNVDSEALVITVTLGERIPALBATADQGAASEMSQSG 353
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 339 FSAATISNGLPLSTVSGSYDAGGVLYTSKE-----KSTF-INMTRVSDMDAYLGYAA 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 354 FSAHYSDVMVLGAVGAYDNGTVVMOKASQIIIPRNTTFNVESIKONEPL-ASYLGYTV 412
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 393 AILNRVOSL-VLGAPOHIGLVAMPFRONTGMESNANVKGTOIGAYRGSLCSVDVD 451
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 413 NSATASSGDLVYTAGQPRYHNTGQVIYRMEDGNTKILQTSGEQISGYFSGILITTDID 472
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 452 SNGSTDVLIGAPHY-----YEQTR--GGQVSCVPLPRGORARWQCDVLY 495
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 473 KDSNTDILLVGAPEWYMGTEKEBGKVYVYALNQTRFYOQMSLEPIKQTCSSRQNSCTT 532
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 496 GEGQGPWG-RPGALVLDVNGDKLTVAGAPGEEDNRGAVYLFHGTSGSGISPSRSQ 554
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 533 ENKNEPCGARFGTAIAAVKDLALDGFNDIVIGAPLEDHGGAVIYHG-SGRTIRKEYAQ 591
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 555 RIAGSKLSPLQVFGOSLGGGLTMDGLVLTGVAQGHVLLRSQFVLVRKAIMFNP 614
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 592 RIPSQGGDKTLKFGQSIHGEWDLNGDGLTDVTIIGLGAALFWRSDVAVKVTNNFBN 651
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 615 EVARNVFECDQVVKGEAG--EVRVCLHVQ-KSTRDLRREGQIQSVVYTDIALDSGRPH 671
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 652 KVMIOKKNCH--MEKERTVCINATVCFEVKLSKEDTIYEADLQ---YRVLDSLRLQ 704
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 672 SRAVFNET-----KNSRROQVGLTQTCETLKLQPLNCLIEDPSPVILNFSLVGT 725
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 705 SRSFFSGQERKQVRNITVRKSEC-----TKHSFYMLOKHDFQDSVR-----ITLDFNLT-D 755
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 726 PLISAFGNLRPLAEDARLFTALFPFEKNGNDNICQDLSITFSFMSLDCLVVGGRPE- 784
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 756 PENG-----PVLDDSLPNSVHEYIPPAKCGNKEKICISLDSLRVATTEKDLIVRSQMDK 810
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 785 FNVTVTRNDGDSYSTQVTFPPLDLSYRKVSTLQNRQSRSLRACESASSTEVSGAL 844
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 811 FNVSLTVKNTKDSAYNTRTVIHYSPNLVFSGIEATQKD-----SCSN----- 853
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 845 KSTSCSINHIPIPPENSEVTFNITFDVDSKASLGN-KLLKANKVTSNNMPTNKTEFQLE 903
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 8

145914

Integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003

C:Accession: I45914

R:Kanata, T.; Fuzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integrin

A:Reference number: A54402; MUID:94193647; PMID:7511592

A:Accession: I45914

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <RAM>

A:Cross-references: GB:125886; NID:9439695; PIDN:AAB59255.1; PID:9439696

F:161-336/Domain: von Willebrand factor type A repeat homology <WMA2>

Query Match 18.5%; Score 1085; DB 2; Length 1170;

Best Local Similarity 27.6%; Pred. No. 7.2e-67;

Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

```

QY 1 FNLDTENAMTQ-ENARGFGQSVVQL---QGSRVVVGAPQEIIVANQSGSYQC--DYST 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 19 YNVGLPKAKI FSGFSSEQFGVAVQQFINPKGNWLLVGSFSPWSPKPNRMDGVYKCPVDLST 78
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 GSCPIRLQ-----VPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENYVRKGLC 107
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 79 TTCSEKLNLTSTMSNVTMKNMNSLGLTLTRNVGTGGFLTCGPLWAOCCGQSYTTGVC 138
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 108 FLFGSNLRQOQKPFPEARLGCQSDSDIAFLVDCSGSIIIPHDFFRAKEFTSTWELQK-- 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 139 SDVSPDP-QLRTSAPAVQTCF-SFIDVVVVVCDSESNIIYPD--AVKNFLEKFVQGLDIG 194
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 KSKTLFSLMOYSEBEFIHFTKEFQNNPNRSLIKPIQTLI-----GRTHATGIRKVVRE 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 195 PTKQMGILQIYANNPRVFNLTNFKSD---EMIKATSIQTQYGGDLTNTKAIQYARDT 251
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 222 LFNITNGARKNAKPLILITDGEKFGDPLGYEDVIPEDRGVIRYVIGV-----GDAPR 276
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 252 AYSTRAGRPQATKVMVVTVDGESH-DGSKLKAVIDCQKNILRFGI AVLGLYNRLD 310
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 SEKSRQELNTVASKPRPDHVFQINNPEALKTIONLREKIPAEIGTGTGSSSEHEMSQ 336
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 311 TKNLIKETKATASIPTERHFNVDSEADLLEKAGTIGQIFSIETVQG-GDNFQOMESQ 369
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 337 EGPSNAIT--SNGPLSTVSGVDNAGGVFLYTSKEKSTFINMT--RVSDNMN-DAYLGYA 391
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 370 VGFSABYSPQNNILMLGAVGAYDWSGTVVQKTPHGLIFSKQAFQEIQLQDRNHSYLSYS 429
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Qy	392	AAIILNRVQSLVGLGAPYOHIGLVAMFRONTGMWESNANV-----KGTOIGAYFGASL	445
Db	430	VASISTGNSVHFVAGAPRANYTOQIVLYSVN-----ENGWTVIQSGRQDQIGSIFGSLV	484
Qy	446	CSVDVDSNGSTDLVLGAPHYBQTR--GGQSVSCPLPRQGRARWQCDVAVLYGEOQPGW	503
Db	485	CAVDVNDKDTITDVLVAGAPYMDLKKRGRVYLFITKG--ILNMH--OFLEGPNGLENA	541
Qy	504	RFGAALTVGLDVNGDKLTDVAIGAPBEDNRGAVLPHGTSGSGISPSHSORITAGS--KL	561
Db	542	RFGSAIAALSDIINDGFNDVIVGSPLENQNSGAVIYNGHEG--ILRYSQKILGSDRAF	600
Qy	562	SPRLQYFGSLGSGQBLTMDGLVDLTVGAGQHVLLRSQPLVRKVAIMEFNPVARNVP	621
Db	601	SSHLQYFGRSLDGLGDSITDVSAGAFQVQVQLWSQIADSVSDAFTPKKI--TL	658
Qy	622	ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALD-----SGRPHSAVEN	677
Db	659	NKVAET-----KLKLP-----SAKPRPTNQNNQVAITNITIDBDQSSSRVIRGLFK	707
Qy	678	ETKNSRTRQTVLGLTQTCB--TLKQLPNCIEDPVSPIVLRINFSL--VGTPLSAFCNL	733
Db	708	ENNERCLQKTMIVSQORCEYIIHQEPS--DIISPLNLCWNISLENGT-----	756
Qy	734	RPVLADDAQRLFTALPPFPKKNCGNDNICODDLSITP-----SFMSLDCLVVGGRPRENVT	789
Db	757	NPALAEYSETVTVFSPHFKDGDGVCISDLVNVQQLPATQQQPFIVSNQKRLTFSV	816
Qy	790	TVENDGEDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESAST--EYSGALKSTS	848
Db	817	QLKXKESAVNTIIVDPSENLF-----ASMSMPVDGTGVTCQIASSQKSVT	864
Qy	849	CSINHPIPPENSEVTNITPDVDSKASLGKLLKLLKANVTSENMPRTNKTBSQLELPVKY	908
Db	865	CNVGYPALKSKQVTFPTNFDFNLQ--NLQQAISIFRALSESQENNAADNSVNLKLSLY	923
Qy	909	AVYVWVTSHGVSSTKYNLFTASENTSRVMOHQVQVSNLQOR-----SLPISLVPLV	958
Db	924	DAEIHIT--RSTNINFTEVSLDGNVSVV--HSFE--DIGKPIKSVKVTGSPVPSMA---	976
Qy	959	PVRANQTVIWDREPQVTPSEN---LSSTCHTKB-----RLPSHSDFLAE--	998
Db	977	-----SVIHIPOYTKXNPLMYLTGVHTDQAGDISCEABEINPLKIGQTSSSVSFSEN	1030
Qy	999	LRKAPVNVCSIAVCQIQCDIPFGIOEBEENATLXGNLFDWYIKTSHNELHIVSTAEI--	1057
Db	1031	FRHKLNCRTASCNIMCWLRLQVKGEYFLNVSTRIMNGTFAASTQTQVLTAAABID	1090
Qy	1058	LFNDSVFTL-----LPGQAFVRSQETETKVBEPFE--VPNPLPLIVGSSVGGLLALITA	1110
Db	1091	TYNPQIVIEENTVTP-----LITMKPHEKVEVPTGVIVGSVIAGILLALLAIVA	1140
Qy	1111	ALYKLGFFKRYKQDM	1125
Db	1141	ILWKLGFPRKRYKDM	1155

RESULT 9
S44142
VLA-2 protein homolog - mouse
C/Species: Mus musculus (house mouse)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
C/Accession: S44142
R/Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A/Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A/Reference number: S44142
A/Accession: S44142
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1178 <EDE>
A/Cross-references: EMBL:Z29987; NID:q473098; PIDN:CAA8287.1; PID:q473099
P:169-344/Domain: von Willebrand factor type A repeat homology <WAA2>

Query Match	18.2%	Score	1071	DB 2	Length	1178			
Best Local Similarity	27.9%	Pred. No.	6.9e-66						
Matches	342	Conservative	208	Mismatches	467	Indels	188	Gaps	44
Qy	1	FNLDTEAMTPO--ENARFGOSVVLQ----	QSRVVVVGAPQBIIVANQRGSLVQC--DYST	54					
Db	27	YVGLPGAKIPSGPSSEQGYVQQLTNPQGNWLLVSPSGFPENRMGDVYKCPVDLPT	86						
Qy	55	GSCPIRLQ-----VPVBAVNSLGLSLAATTSPQQLLACGPTVHQTCSNTYVKGJL	107						
Db	87	ATCEKLNQNSASISNTVEIKTNMSLGLTTRNPTGGTGLCGPLWAHQCNQYATGIC	146						
Qy	108	FLFGSLNRQQPQ---KPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAXEPFISTVMEOL	164						
Db	147	---SDVSPDFQPLTSPSPAQCFSL--VDVVVWCDESNSIYP--WEAVKQFLVKFVFTGL	199						
Qy	165	K--KSKTILPSLMQVSEBFRHFTFKFQNNPNRSLIKPITCLLG--RHTATGIRKVVRE	221						
Db	200	DIGPKKTOVALIQVANGPRIIFNLNDFETKEDMVOATSETQHGGDLTNTFAIEPARY	259						
Qy	222	LFNITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAPR	276						
Db	260	AYSOTSGRPGATKVMVVVTDGESH--DGSKLKTVIQCCNDDEILRFGIAVLGYLNRALD	318						
Qy	277	SEKSRQELNVTASPPRDRHVFQINNFEALKTIONQLREKIPIABGTQSGSSSFHEMSQ	336						
Db	319	TKNLIKEIKAIATSTERYFFNVADAEALLEKAGTLGEQIFSIETGVQ--GDNQWEMAQ	377						
Qy	337	EGFSA--AITNSGPLLSTVGSYDWAGGVFLVTSKSTFINMT--RVSDMN--DAYLGYA	391						
Db	378	VGFSADYAPQNDILMLGAVGAFDMSGTLVQETSHKPVFPKQAPQVLDQRNHSFLQVS	437						
Qy	392	AAIILNRVQSLVGLGAPYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCS	447						
Db	438	VAAISTEDGVHVFAGAPRANTYTGQIVLYSVNK--QGNVTVIQSHRGDQIGSYFGSVLCS	494						
Qy	448	VVDVDSNGSTDVLIGAPHYBQTR--GGQSVSCPLPRQGRARWQCDVAVLYGEOQPGWRF	505						
Db	495	VVDVDTITDVLVAGAPYTMNDLKKRGRVYLFITKGILNCHQ---FLEGEPTGNARF	551						
Qy	506	GAALTVGLDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHSORITAGSKLSR--	564						
Db	552	GSAIAALSDIINDGFNDVIVGSPVENQNSGAVIYNGHQGT--IRTKYSQKILGSGNAPR	610						
Qy	565	LQYFGOSLGGQDLTMDGLVDLTVGAGHVLLRSQPLVRKVAIMEFNPVARNVPFEC	623						
Db	611	HLQFFGRSLDGLGDSITDVSIGALQVQLWSQIADVAIEALFTP-----	660						
Qy	624	NDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALD---SGRPHSAVFNET	679						
Db	661	-DKITLLNKDAKITLCLCFRAEFPAGQNNQV--AILFNMTILDAGHSSRVTSGVFRFN	717						
Qy	680	KNSTRQTVLGLTQTCET--LKQLPNCIEDPVSPIVLRINFSLVGLTGLSAFGLRPLV	737						
Db	718	SERFLQKNVNVNEVKCEHHISIQKS---DVVNPLDLRVDISLENPGTS-----PAL	768						
Qy	738	ABDAQRLFTALPPEKKNCGNDNICODDLSI-----TFSPMSLDCLVVGGRPREFNT	788						
Db	769	EAYSETVTVKVSIPFYKCEGSDGICISDLILDVQQLPAICTQSF-----IVSNQKRLTFS	823						
Qy	789	VTVRNDGEDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESAST--EVSGLAKST	847						
Db	824	VILKRGESAVNTVVLAEFSENLF-----ASFSPMPVDGTVEVTCVSGSQSV	871						
Qy	848	SCSINHPIPPENSEVTNITPDVDSKASLGKLLKLLKANVTSENMPRTNKTBS--FQLELP	905						
Db	872	TCDVGYPAKSECOVTFITNDFNLQ--NLQQAIFNFOAFSESQ--ETNKADNSVSLTIP	928						
Qy	906	VKAVYVWVTSHGVSSTKYNLFTASENTSRVMOHQVQVSNLQOR-----SLPISLV	955						
Db	929	LLYDAELHLT--RSTNINFTEVISESDENAPSVIK---SVEDIGPKFIPLSKVTAGSAPVMA	984						

QY 956 FLV-----PVLNQIVWDRQVTF-SENLS 980
D 985 LVTHIPQYTKENPLLYLTGIOTDAGDISCTAEINPLKLPHTA-----PSVSFKENFR 1040
QY 981 STCHKERLPSSHDFLABRKAPVNGSLAVCQIQDIDPFQIGESFNATLKNLSFDW 1040
D 1041 ---HTKE-----LDCTTSCSNITCWLKDLHMKAEYFNVTRVWRT 1086
QY 1041 YIKTSHNHLIVSTAEILFNDVSTLPLGQGAFAVRSOTETKVRPFVNPPLPLIVGSVG 1100
D 1081 FAASTFTQVLTAAAEIDTNPOLFVIEENAVTILMIMKPKAEVPT--GVIIGSIIA 1138
QY 1101 GILLALLITATLYKLGPPKROKDM 1125
D 1139 GILLALLATAGLWKLGGPPKROKDM 1163

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAK>
A:Cross-references: GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
B:Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIc)
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, v. 18-21 <ZUT>
A:Cross-references: GB:L24121; NID:G400342; PIDN:AAA16619.2; PID:G4583535
A:Note: authors translated the codon GGA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1068; DB 2; Length 1181;
Best Local Similarity 26.9%; Pred. No. 1.1e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFQ-ENARFGQSWQL--QGSRVVVGAPQIRVAAQRGSLYQC--DYST 54
D 30 YNVGLPEAKIFGSPSEQGYAVQOFINPKGNLLVSGSPWSPENRNGDVVKCFVDLST 89
QY 55 GSCPEIRLO-----VPVEAVNMLGLSLAATTSFPOLLACGPTVHQTCSNTYKGLC 107
D 90 ATCEKLNLTSTSIPIPVNTEKTNLSGLILTRNMTGGLTCGPLWAOCCNQYITTGVC 149

QY 108 FLGSLNRQOPKPPALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK-- 165
D 150 SDISDPF-QLSASPSFATPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKVFQGLDIG 205
QY 166 KSXTLSLAQYSSEFRIHFTFKFQFNPNRSLIK2ITQLLG-RTHATGIRKVRLEFN 224
D 206 PTKQVGLIQYANNPRVFNLTNTKBEIVATSTQSVQVGGDLINTFGAIVQARYAYS 265
QY 225 ITGARKNAPKILILTDGEKFDGPDGYEDVPEADREGVIRVIGV-----GDAFRSEK 279
D 266 AASGGRSRTKVMVVVVDGESH-DGSMLKAVIDOCNHDNILRFGIAVLGYLNRLDITKN 324
QY 280 SROELNTVASKPPEDHVFQINNFEALKTIQNLREKIFAIEGTQTCTGSSSFEHMSQEGF 339
D 325 LIKEIKAIASIPTEYFFNVSDAEALLERKAGTIGEQIFSEGTVOG-GDNFQEMSQVGF 383
QY 340 SAAITNGP--LLSTVGSYDWAGEVFLYTSKEKSTFINMT--RYVSDMN-DAYLGIAAAL 394
D 384 SADYSSONDILMAGVAGFAGSGTIVTKTSHGHLIFPKQAFDQILQDRNHSSVLYGSVAA 443
QY 395 ILNRVQSLVLAGPRYCHIGLVAKFQNTGMWESNAV-----KGTQIGAYTGAISCSV 448
D 444 ISTGESTHVFAGAPRANYTQIIVLYSVN-----ENGNIIVIAHAGDQIGSYFGLCSV 498
QY 449 DVDSNGSTDLVLIGAPHYEYQTR--GQOVSVCPPLPRQORARWOCDAVLYGEQOPWGRFG 506
D 499 DVDKDTITDVLVAGAPWMSDLKKEGRVYLFITKGIILGQH--FLEGEPIENTRFG 555
QY 507 AALTVDVNGDKLTVAIGAPCEENRGAVLFGHTSGSGISPSHSORTAGS--KLSPR 564
D 556 SAATAALSDIIMDGFNDVIGSPLENQNSGAVIYINGHQT-IRTKYSQKILGSDGAFESH 614
QY 565 LQYFGSLSGGQDLTKMDGLVLTGAGHYLLRSQVLRVKATMEPNPREVARNVEFN 624
D 615 LQYFGRLDGYGLNGDSITDVSIGAFQVQVQVWSQSADVAIEASTPEKI--TLVNKN 672
QY 625 DQVYKGEAGEVRVCLVQKSTDRLEBGOIQSVVYDLDALD---SGRPHSAVFNVEK 680
D 673 AQII-----LKLCP-----SAKPRPTKQNNQVAIVNITLDDAGFSRVSRLFKENN 721
QY 681 NSTRRTQVLGLTQTC--ETLKLQPNCEIDPSPVLRNLFSLVNGTPLSAFGLNLRPVL 738
D 722 ERCLQKNVVNQASCPHEIYIQES--DVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALPPEKNGCNDNIQDDLSITF-----SPMSLDCLVVGQPRFNVTVTVRND 794
D 773 AYSETAKVFSIPFHKDCGEDGLCISDLVDVRQIPAAQEQPFIVSNQNKELTFSTVLNKK 832
QY 795 GEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESAST-EVSGALKSTSCSINH 853
D 833 RESAYNTGIIVDFSENLPF-----ASFSLPVDGTEVTCQVAAASQKSVACDVGY 880
QY 854 PIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEFQLELPVKYATMV 913
D 881 PALKREQQVTFITNFDNLQ-NLQNASLSFQALSSESQEBENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSHGVSITKYNFTASENTSRVMOHQVQVSNLQOR-----SLPISLVFIV----- 958
D 939 ---HLTSTNINFEISSDGNVFSIVHSFEDVGPKTFSLKVTGTVSVVNAVVIHIIPQ 995
QY 959 -----EVLNQTVIWDPRQVTF-SBNLSSSTCHTKER 988
D 996 YTKENPLMVLTVGTQTKAGDISCNADINPLKIGQT-----SSSVSPKSENFPR---HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNVCSIANQRIQCQDIPFGIQEBENATLKNLSFDWYIKTSHNH 1048
D 1048 -----LNCRTASCNVTCWLKDVHMKGEFVNVTTIRWNGTFPASSTFQT 1091
QY 1049 LLIVSTAEI-LFNDVSFTLLFGQAFVRSQTEKVPFEPFNP-----LP--LIVGSSVG 1100
D 1092 VQLTAAAEINTYNPPIVVI-----BDNTVTIFLMIMKPKAEVPTGVIGSIIA 1141

QY 1101 GLILLALITRALYKLGFFKQYKDM 1125
 DB 1142 GILLILALVAILKLGFFKQYKDM 1166

RESULT 11
 A35854
 Integrin alpha-1 chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 15-Sep-2003
 C:Accession: A35854; S11243
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, P.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
 A:Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin
 A:Reference number: A35854; MUID:90338125; PMID:2380249
 A:Accession: A35854
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1180 <IGN>
 A:Cross-references: GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
 P:170-345/Domain: von Willebrand factor type A repeat homology <WMA2>

Query Match 18.1%; Score 1062; DB 2; Length 1180;
 Best Local Similarity 27.4%; Pred. No. 2.9e-65;
 Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

QY 1 FNLDTENAMTFQENARG-FCQSVVQL---CGSRVWVGAPQEIIVAAQNGSLYQCYSTGS 56
 DB 29 FNVDKNSMFGSPVEDMFGYTVQYENEGKVLGSPVLPQPKARTGDKVYKCPVGR 88
 QY 57 CEP-IRLQVPVEA-----VNMSLGLSLAATSPQQLACGTPVHQCSENTRYKGL 106
 DB 89 AMPCKLDLPVNTSIPNVTBIKENMTFGSTL-VTPNPGFLACGPLYAYRCGLHLYTTGI 147
 QY 107 CFLFGSNLRQOPKQPEALAGCQEDSDIAFLVDGSGSIIPHDFRAKEPISTVMQLK- 165
 DB 148 CSDVSPTFQVNSNAP--VQECSTQ-LDITVLDGSGNSIYP--WESVIAFLNLDLKMDOI 202
 QY 166 -KSKTLFSLMOYSEPRIFHTFEFQNNPNRSLIKPIQLLG-RHTATGIRKVRREL 223
 DB 203 GPKQTQVGIQVGENVTHEFNANKYSTEEVLVAANKIGRQGLQTLTALGIDTARKEP 262
 QY 224 NITWGARKNAPKILITLDEKFGDPLGYREDVPEADREGVIRYVIGVDAPR-----SE 278
 DB 263 TEARGARGVKVMVIVTDGESH-DNYRLKQVTDQCDENIQFSAIALGHYRGNLSTE 321
 QY 279 KSRQELMTVASKPRDHVFOINNPEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEG 338
 DB 322 KVEEIKSIASEPTEKFEFFNVSDALVTVKALGERIFALEATADQSAASFENMSQTG 381
 QY 339 PSAAITNGPLLSVGVNAGGVFLYTSKEKSTFIMNT--RVDSMDND---AVLGYAAA 393
 DB 382 FSAHYSQDWMVLMGAVGAYDNGTVVMQKAMQVPIHNTTFTQTEPAKNEPLASLYGTYN 441
 QY 394 IILNRRVQSLVLAGPRYOHILVAMPQNTGMSWESNANVKGTOIGAYFGASLCSVDVDSN 453
 DB 442 SATIPGDVLYTAGQPRNHTQVVIYKXEDGNINILQTLGGEQIGSVFGSVLTIIDDKD 501
 QY 454 GSTDLVLIGAPHY-----YDQTR-GGQVSVCPILPRGARWQCDVLYGE 497
 DB 502 SYTDLLLVGAPWYMGTEKEQGVVYVAVNQTRPEYQMSLEPIRQTCSSILKONSCTKEN 561
 QY 498 QGQPWG-RFGALTVLGDVNGKLTVAIGAPGEDNREGAVLYPHGSGSGISPHSQRI 556
 DB 562 KNEPCGARFGTAIAAVXDLNDVGDENDVWIGAPLDDHAGAVIYHSG-KGTIREAQAQRI 620
 QY 557 AGSKLSRLQVFGQSLGGQDLTWDGLVLTVAQGHVLLIRSQPVLKAIKFEKNPREV 616
 DB 621 PSGGDKTLKFFGQSIHGEDMLNGDGLTDVTIGLGGAAALFWARDVAVAVVWVNNFENKV 680
 QY 617 ARNVFECNDQVVKKEAG--ZVRVCLHVO-KSTFDRLREGQISVWTVYDLALDSGRPSR 673

DB 681 NIQKNCER---VEGKETVCINATWCFFHVKLAKEDSIYEADLQ-----YRVTLDSLRLQISR 733
 QY 674 AVFNET-----KNSSTRQTOVLGLTQTCETYLKQLPNCI-----EPVSPVILRL 718
 DB 734 SFFSGTQERKIQRIITVRES-----CIRHSFYMLDKHDFQDSVRVTL 776
 QY 719 NFSLVGTPLSAFAGNLRPVLAEDAQRLFTALFPFFKNGCNDNICODDLSTITTFSMELDCLV 778
 DB 777 DNLUT-DEPENG-----PVLDDALPNSVHEHLPFAKDCGKERCISDLTLNVTSTTEKSLLI 830
 QY 779 VGGPRE-FNVTVTVNDGEDSYRTQVTFPFFLDLSYRKVSTLQNSORSORWELACESASS 837
 DB 831 VKSQCHKFNVSILTVKNGDSAYNTRTVVQHSNPLIFSGIEBKQD-----SCSEN-- 880
 QY 838 TEVSGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLL-LKANVTSENNMPTN 896
 DB 881 -----QNTICRVGYFPFRAGETVTFKLIIFQFNTSHLSENAIHLSATSEEPLESIN 933
 QY 897 KTEFOLEPLVKYAV---YMWVTSHGVT-----KYLNTFASNTSRVMQHOQVSNL 945
 DB 934 DNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPPEINST--EDIGNENINVFYTIKRR 991
 QY 946 GQRLP---ISLVP-----LVPVRLNQTIVWD-----RP-----Q 972
 DB 992 GHFPMPELQLSISFPNLTAQGVPLVPIG-----WSSSDNVNCRPSLEDPFGINSKK 1045
 QY 973 VTFS-----ENLSSTCHTKERLFSHSDFLAELRKAPVNVCSIAVCQRIQCDI--- 1019
 DB 1046 MTISKSEVLKRTGIDQCSSTC-----GVATITCSLLPSDLSQVNVSL 1088
 QY 1020 ---PFGIQEEP---NATLKGNLSDWYKTSNHLILIVSTAEILFNDSVFTLLPQCGAF 1073
 DB 1089 LMKPTF-IRAHFSSLNLTLRGELK-----SENSLTSSSN----- 1123
 QY 1074 VRSQTEKVEPEVNPPLPL--IVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128
 DB 1124 RXBELAIQISKDLGPRVPLWILLASAPAGLLLLMLLILALWKIGFFKPLKKQMEK 1180

RESULT 12
 A41131
 Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
 N:Alternate names: Integrin alpha-4
 C:Species: Mus musculus (house mouse)
 C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
 R:Accession: A41131; S16742
 R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
 J. Cell Biol. 115, 1149-1158, 1991
 A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
 A:Reference number: A41131; MUID:92064645; PMID:1840602
 A:Accession: A41131
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1039 <NEU>
 A:Cross-references: EXBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51485
 C:Superfamily: Integrin alpha-4 chain
 C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;
 Best Local Similarity 22.7%; Pred. No. 1.5e-37;
 Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGR--VVGAPQEIIVAAAN---QRGSLYQCY 52
 DB 41 YMLDPENALLYQGPSOTLFGYSVWLHSHGSKWLIIVGAPTASWLSNASVNVPCAIYRCGI 100
 QY 53 STG---SCPIRLQVP-----VEAVMSLGLSLAATSP-PQLAOG---PTVHQ 95
 DB 101 RKNPNQTCQLQSGSPGSPGCKTCLIEDNQLWGLVTLRSQPGENGSIYTCGHRWKNIFY 160
 QY 96 TSENIVYKGLCFGLGNSLRQOPKQFPPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKE 155
 DB 161 MRSNDKPLPGICVWPSDLRLTELK----- 185

156	QY	P1STVMEQLKSKTLP1SLMQVSEBFRHFTFKPEFQNNPNPSLKIPTQLLGRTHHTATGI	215
186	DB	-----RMAPCYKDT-----	195
216	QY	RKVRELFNITNGARKNAFKILILITDKEKGGDPLGYEDVIPZADREGVIRVIGVGDAF	275
196	DB	-----RKGE-----	200
276	QY	RSEKRSQBSNTVASKPPRDHVFQINNFRALKXTIQNOLREKIPAEICTQTGSSSFEHMS	335
201	DB	-----NFAS-----	205
336	QY	QBGFSAAITNSGPLLSTVGSYDMAGGVPLY--TSKEKSTFINMTVRVDSMDNDAYLGYA--	391
206	DB	QAGTSSFTQDLIVMGAPGSSWYCTVPVYNITNQYKAFVD--RQNVKSGSYLGSVVG	263
392	QY	AAAILRNVSQSLVIGAPRYQHIGLVAMPONTGWMESNAV----KGTQIGAYFGASLCS	447
264	DB	AGHFSAPHTTEVVGAPQHEQIGKAYIP----SIDENELNIYVEMKKGKLSYFGASVCA	319
448	QY	VDVDSNGSTDLVIGAPHYVEQTRGVQSVCPLRGQBA--RWQCDAVLYGQGPWGRFG	506
320	DB	VDLNDAGFSDL--LVGAPMQSTIRESGRFVY--INSGMGAVVMEMERVLVGDXYA--ARFG	376
507	QY	AALTVLGVGVNGDKLTVAIGAPGEDNRCNAVTLFHTSGSGISPSHSORIASGLSPRLQ	566
377	DB	ESIANLGDINDGFDIAIGAPQEDLRGAVIYNGRV--DGSSITSYQRIEQQIKSLR	435
567	QY	YEGQSISGGODLTMDGLVDLTVGA--QGHVLLRSQPVLVKAIMEFNPREVARNVFCN	624
436	DB	MEGQSISGQIDANNNGYVDVAVGAPQSDSAVLLRTPVVIVEASLG--HPESVNRKFPDCT	494
625	QY	DQWKGKEAGEVRVCLHVOKSTRDLRREQIOSVVTYDLADSGR----PHSAVP--NET	679
495	DB	-----ENGLPSVNCMLTLCFSYKGEKVPGYIVLPYNVSLDVHRKASPSRPFYPSNGT	547
680	QY	KNSTRQTOVLGTOFCETLKLQLPNCIEDPVSPIVLRINFSL-----VGTPLSAFGNLR	734
548	DB	SDVITSGIRVSSGGEKCRTHQAFMRKQVDRDILTPIHVEATYHLGHVITKRNTEBPPPLQ	607
735	QY	PVLAEDAQR--LFTALPFPKPCNGNDNICODLSITPS-----PMSLDCVLVGGREFNV	787
608	DB	PILQOKKQKDVIRKMTNFARFCAYEN--CSADQVSAKAVGLKPYENKTVLAVGSMKTHL	666
788	QY	TVTVRNDGSDSYRTQVTFPPFDLSYRKVSTTLQNORSQRWSRLACBASSTVEVGALKST	847
667	DB	NVSLFNAGDDAYETLLNVQLPTGLYFIKULDEEK-----QINCE--VTESSGIVK--L	716
848	QY	SCSINPIHPENSEVTNITFDVDSASLGNKLLKANVTSEN--NMPRTNKTETFOLELP	905
717	DB	ACSLGYIYDRLSRIDISFLDLVSSUSRAHEDUSISVHASEGEGDQVRDNRTVLTTP	776
906	QY	VKVAVYVAVTVSHGV--STKVINFTASNTSRVMQHQ-----YQVSNLGRSLP--ISLVFL	957
777	DB	LRFEV--MLTVHGLVNPSTS FVYSGSSENEPETCMAEKLMLTPHVTIGISMAPNVSVKIM	834
958	QY	VPVRLNQTVIWRDP--QVTFSENLSSTCHKE-----RLPSHSDFLAEALR	1000
835	DB	VP---NSFLPQDDKLFNVLVDQVTTGQCHPKYHRECTPAQQKGIAGTLTDDIVKFLSKTD	891
1001	QY	KAPVWNC5IA--VCQRITQCDIPFGIQEBFNATLKGNLSPDWYIKTSHNHLILIVSTAEL	1058
892	DB	KR--LLYCWKADQHCCLDFLN--FGWMESG-----KEASVHIQLEGRPSIL	933
1059	QY	FND5VFTLLPGQAGVAVRSQVETKVEPEVNP-----	1090
934	DB	EMDETSSL-----KPEIKATAPPEPHPKVIELKNDENVAVFLGLHQRPKRHF	983
1091	QY	LPLIVSSVCGHLLALITAAALYKLGFFKRYQNDWASE	1128
984	DB	TTTTTSTLGLVIAHIAISCWMMKAGFFKRYKYSILQE	1022

RESULT 13

S06046

integrin alpha-4 chain precursor - human

N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain

C;Species: Homo sapiens (man)

C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999

C;Accession: S06046; A39355; D28018

R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.

EMBO J. 8, 1361-1368, 1989

A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other integrin

A;Reference number: S06046; MUID:89356603; PMID:2788572

A;Accession: S06046

A;Molecule type: mRNA

A;Residues: 1-1038 <TA>

A;Cross-references: GB:X16983; EMBL:X15356; NID:G33945; PIDN:CAA34852.1; PID:G33946

R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.

Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991

A;Title: Characterization of the alpha-4 integrin gene promoter.

A;Reference number: A39355; MUID:91239513; PMID:2034655

A;Accession: A39355

A;Molecule type: DNA

A;Residues: 1-93 <ROS>

A;Cross-references: GB:M62841

R;Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A;Title: The very late antigen family of heterodimers is part of a superfamily of molecules

A;Reference number: A94151; MUID:87204112; PMID:3033641

A;Accession: D28018

A;Molecule type: protein

A;Residues: 40-50, 'E', 52-53 <TA2>

C;Genetics:

A;Gene: GDB:ITGA4; CD49D

A;Cross-references: GDB:I128032; ONIM:I192975

A;Map position: 2q31-2q32

C;Superfamily: integrin alpha-4 chain

C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmembrane

E:1-39/Domain: signal sequence #status predicted <SIG>

E:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 635; DB 2; Length 1038;

Best Local Similarity 25.3%; Pred. No. 1.3e-35;

Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAPRSEKSRQELNTVASKPRD-----HFVQINNFEALKT-----IQN 310

DB 121 GKTCLERDQNLGVTLSRQFGENGSIVTCGRHWKNIIFYIKENKLPGTGCYGVPPDLRT 180

QY 311 QLRKI-----FATEGTQSSSFEHMSQEGPSAAITNSGPIILSTVGSY 356

DB 181 ELSKRIPACYQVYVKFGENFA-----SCQAGISSFTKDLIVNGAPGSS 225

QY 357 DWAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGVA--AAIILNRVQSLVLAGAPRY 411

DB 226 YWTGSLFVYNIITNKYKAFDKDQNVKF---GSVLGYSVCGAGHFRSQHTTEVVGGAPQHE 282

QY 412 HIGLVAMFRONTGWESENANV---KQTQIAYFGASLCSDVDSNGSTDLVLIGAPHYY 467

DB 283 QIGKAYIF---SIDEKELNILEMKKGLGSYFGASVCAVDLNAUGFSDL-LVGAQWQS 337

QY 468 EQTRGGGVSVCPILRGQRARWQC-DAVLYGEQGFMRFGAALTVLGDVNGDKLTDVAIG 526

DB 338 TIREGRVFVY-INSGSAVNMVETNLVGSDKYA-ARFGESIIVNLGDIIDNGFEDVAIG 395

QY 526 EQTRGGGVSVCPILRGQRARWQC-DAVLYGEQGFMRFGAALTVLGDVNGDKLTDVAIG 526

DB 338 TIREGRVFVY-INSGSAVNMVETNLVGSDKYA-ARFGESIIVNLGDIIDNGFEDVAIG 395

QY 527 APGBEDNRGAVYLFHGTSGGISPSHSQRTAGSKLSRLOYPFGQSLGGQDLTMDGLVDL 586

DB 396 AFQEDDLQGAIIYINGRA-DGISSTFSQRIEGLQISKLSNFGQISGQIDADNNGVYDV 454

QY 587 TVGA--QGHVILLRSQPLRVKATMEFNPREVARNVEPCNDQVVKGEAGEVRVCLHVQK 644

DB 455 AVGAFRSDASVALLTRPVPWIVDASLS-HPBSVNETKDC-----VENGMFSCIDLTL 506

QY 645 STRDLRIEQQI-OSVWYVDYDALDSGR-----PHSRVAFNETKNSTRQTVLGLTQTCETL 699

507 CFSYKGVGVYVLPYNNSLDNRKAEPPRPYSSNGTSDVITGSIQVSSREANCRT 566
700 KIQLPNCIEDPVSPIVLRNFSLVGTPLS-----AFGNLRPVLAEDAQR-LFTALPPEK 753
567 QAFMRXDVRLDILPTIQIEAAYLGHPIVSKRSTEBPPLQILOQKXKDIKKKTINFA 626
754 NCQNDNIQDQLSIT--FSGM-----SLDCLVGGPREFNVTVRNDGSDSYRTQVTFPP 807
627 PCAHEN-CAADLQVSAKIGFLPKHENKTYLAVGSMKTLMLNLSLNFAGDDAYETTLHVKL 685
808 PDLVSRKVTSTIQNQRORSRLACBSASTEVSGALKSTCSINHIPIPPENSEVTFNIT 867
686 PVGLYFIKILEEEK-----QINCE---VTDNSGVVQ-LDCSIGIYYVDHLSDIDISFL 735
868 FVDVSKASLGNKLLKANVTSEN--NMPRTKTEFOLEPLPVKYAVVMVTSHGVTGKVLN 925
736 LDVSSLRAEEDLSIVHATCENEBEEDNKKSRVTVAIPLKYEVKLTVHGFVNTSFY 795
926 FTASENTSRY-----MHOYQVSNLQGRSLP-ISLVFLPVVRLNQTIVWDRPQVTFSEN 979
796 GSDNEPEPTCMVEKGNLTFHVINTGSMAPNVSVBIMVENSFS-----PQTDKLFNI 848
980 -----SSTCHTKERLPSSHDFLAEIRKAPVNVCSIAVCQIQDIPFGIQEBFNATLK 1033
849 LDVQTTTGECHFNQ-----RVCALEQ-----QKSAQOTLK 880
1034 GNLSPDWYIKTNSHLLIIVSTAE-----ILFN-----DSVFTLLPGQAFVRSQFE 1079
881 GIVRF--LSKTDKLLYCIKADPHCLNFCNFGKMSGKEASVHIQLEGRPS-ILEMDE 936
1080 TKVEPPEV-----RNP-----LPLIVGSSVGGLL 1103
937 TSALKEFEIRATGPEPNPRVIELNDENVAHVLEGLHQRPKRYPTIIVSSLLGLI 996
1104 LLALITAAALYKLGFFRKQYKQDMSE 1128
997 VLLLIISYVWVKAGFFRKQYKSIILQE 1021
RESULT 14
I58409
Integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
A:Cross-references: GB:D25303; NID:G464180; PIDN:BA04984.1; PID:G533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:I24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 2.9e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;
QY 331 EHEMSQGGFSAITSNGPLLSITGVSYDWAGVGLYTSKESKSTFINWT-RVDSMDNDAYLG 389

190 EHGSCQAGIAGFTTEELVWVGAPGSPYAGTGIKVLNLTD-NTYLKLNDEVIMRRYTYLG 248
390 YA--AAIILNRVQSLVAGAPRYOHIGLWAFR--QNTGMWESNANVKGTOIGAVFGASL 445
249 YAVTAGHPSHSTTDVGGAPQDGIKGVYIFRADRSGLIKIFOASGKKMGSYFGSSL 308
446 CSVDVDSNGSTDLYLIGAPHYYEQTGQVSVCPPLFGQARMOCDVLYGEOQGPWGRF 505
309 CAVDLNGDGLSDL-LVAGAPMFSEIRDEQVTVY-INRNGGALDEE-QLALATGD-GAYNAHF 364
506 GAALTVLGDVNGDKLTVAIGAPGEEDNRGAIVLPHGTSGSGISPSHSORISAGSKLSPL 565
365 GESLADLNDLNDGFPDVAIGAPREDDPAGAVIYHGDAG-GIYVQISMKLSGQKINPVL 423
566 QYFQGSLSGGQDLTMDGLVDLTGCA--QGHYLLRSQPLVLRVKAIMEFNPPEVARNPFC 623
424 RMFQSISSGIDMDGNGYDVTVGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
624 NDQVVKGEAGEVRY-CLHVQ-----KSTRDLRBEQIQSVVTVYDIALDLSGRPHRAV 675
483 HD-----GQQPVNCLNVTTCSFHGKHVPBEI---GLNYVLMDADVAKKEKQOMPVY 531
676 F--NETKNSTRRTQVIGLGTQTCETLKLQLPNCIEDPVSPIVLRNLFSL-----VGTPLS 728
532 FVLLGETMGQVTEKLQLTMYEETCRHYVAHVRRVQDVISPIVFEAAVSLSEHVGTBEER 591
729 AFGNLRPVL-----ABDAQRLLTALFPFKNGCHNDICQDLSITFSFM--SLD-----CL 777
592 ELPLTTPVLRWKKQKIAQKNQTV---PERNCSED-CAADLQGLQGLLSMDDEKTLYL 647
778 VVGGPREPNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSORSLACESASS 837
648 ALGAVKNISINISINLGDADYDANVSNVSELPF-----INWQKEENGISCLELLES 701
838 TEVSGALKSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRINK 897
702 DPL-----KCSVGFPPMRSKSYKFSVIPDTSLS--GEEVLSPIVTAQSG--NTER 750
898 TE-----FOLELPVKVAVVMVTS-----HGVSTKYLNFNTASENTS---RVWQHGY 940
751 SESLHNDTLVLMVPLMHEVDTSITGIMSPTSFVVGESVDAAFNQLDLDLECHQFQINITL 810
941 QVSNLQGRSLPISLWFL-VPVRLN-----QTVIMDRPQVTFPSENLSSTCHTKER 988
811 QVYNTGPSTLPQSSVSIISFPNRLSSGGAEMFHVQEMVVGQEKNGCSFOKNPTPCIIPOEQ 870
989 LPSHSDPFLAEIRKA-----PVNCSIAVCQRIQCDIPFGIOEBFNATLKGNLSFD 1039
871 ENIFHTIIFATFKSGRKVLDCEKPGISCLTAHCN-----PSLAKESRTI-----D 917
1040 WYIKTSHNHLIIVSTAEILFNDISVFTLLPGQAFVRSQTETKVEP----- 1084
918 IY-----MLNLT-EILKXSSSVIQ-----FMSRAKVQVDPALRAVVEIAHGNPEEV 962
1085 ---FVFPN---PLPLIVG-----SSVGGLLLLALITAAALYKLGSEKRYKQKDM 1126
963 TVVPEALNLEPRGYVGVWGIITAILSLVGLILFLLAVLLWKGPPRRRIKRII 1015
RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA

A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g32202040; PID:g32202041; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 3.5e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAIITSGP--LLSTVGSYDAGGVFLTSKEKSTFINTRVDS-----DMNDAYLG 389
DB 181 QAGFGIIFEDNSALVMGAPGYIYQGIYVOSLLNRSV-VORTQESNTGYTSPDNSTNG 239
QY 390 YAAAI--ILNRVQSIIVLGAPRYQHI-GLVAFPRONTGMWESNANYKGTQIGAYFGASLC 446
DB 240 YSLALGDFNGDGVDYVVGTPRAESLMGLVAIFDQNLQFN--QVMGTQIVAYFGYSVT 296
QY 447 SYVDVNSGTDLVILGAPHYETQRCGVSVCPLEPGQARWQCDVLYGEQ----- 498
DB 297 VVDI--NNDTYDLDLVGAPMYMDGPAIQ-----RWEAGAVYYVLQNPDPVGGA 343
QY 499 -----GQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRCGAVYLFHGTSGS 546
DB 344 SNRLSSTLIGQIIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVIYHG-SAN 402
QY 547 GSPSHSQRIAGSKLS-PRLOVFGSLSGQDLTMGLVDLTVGAQ--GHVLLRSQPVL 603
DB 403 GLKSTPAQVLTPTSLTGHSGITTFPSLQGGQMDKNKYPDLLWGAESANTAVLIRTPVY 462
QY 604 RYKAIMFEPREVARNVFECNDQVVKGEAVRVCLHVOKSTRDLREGQIQSV----- 658
DB 463 SLDATLNTPE-----IGINLENKTYE-LADGTWTSFIAMT 497
QY 659 -----VYVDLALDSG-RPHSRVAFNETKNSRROTQVGL-TQCTETKLQ 702
DB 498 CFTYTGNYLPHDIDISVTVVDSIIANRRAMPVDMDSEITKTRRLAVSTQPCDPLRAY 557
QY 703 LPNCIEDPSPVLRNPLSVLGTPLSAPGN-----LRPVLAEDAQRLPTALPPEK 753
DB 558 VGNSEDKLTPIKVTQYDL-----NDESRLOPHEILPDIIDMATSTQTKQVSION 609
QY 754 NCGNDNIQQDLSITTFMSLDCLVGGPREFNVTVVRNDGSDSYRTQVTFPFLDLSY 813
DB 610 NCYN-NICIFDLDVTVT-PNLNIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLQOF 667
QY 814 RYVSTLQNRQSRWRLACESASSTEVSGALKSTCSINHPIPEX-----SEVTFNIT 867
DB 668 VRL-----ERKANMDPSVTCESDSD-----LRITCDTGNPMVGNKILPEGLTLSTFQVS 717
QY 868 FDVDS-----KASLGNKLLKANVTSNNMPTNKTEFQLELPVKYAVYVNVTSHGVS TK 922
DB 718 GDKDSIEFYFKASENS--EDPNTLENLNMVTVTVDCITLKLNSYPEIWMYSTQED 775
QY 923 YL-----NFTASENTSRVWQHGYQV-----SNLQGRSLPIS-----LVFLVPV 960
DB 776 YVVPFPKAKNASADIGMEVW-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834
QY 961 RLNQTVLWDRPO-----VTESENLSSTCHTKBELPSHSDFLAELRKAPVY 1005
DB 835 MTEEGVTCQLTQKANEGLAPSTAKLSNSTTQVSGRKRREPEVAEALQTDN--VI 892
QY 1006 NCSTAVCQRIQCDDIFFFGIOEEFNAT-----LKNLSFDWYIKTSHNHLIVSTABILF 1059
DB 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY 1060 NDSVFTLLPGGAPVRSCQTKVE--PFEPVNP-----LP----- 1092
DB 929 QKAVSELTPVQVATIASASAANKTIPYNIPLRDFSDSTKASTLVTEELVPPVPIAW 988
QY 1093 -LIVGSSVGGLLLALITAAALKLGFPR-----QYKDMMS--EGGPP 1132

Db 989 WIIVSVLGGIILLIILLGLWKCGPFRKKFGKEKEYAPVASADKGGPP 1038
Search completed: June 7, 2004, 17:18:08
Job time : 20.559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec

Title: US-09-902-481b-5
Perfect score: 5876
Sequence: 1 FNLDTENAWTQENARGFGQ.....FKRQYKDMWSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5846.5	99.5	1152	1 ITAM HUMAN	P11215 homo sapien
2	4470	76.1	1153	1 ITAM MOUSE	P05555 mus musculus
3	3459	59.0	1163	1 ITAX HUMAN	P20702 homo sapien
4	3411	58.0	1162	1 ITAD HUMAN	Q13349 homo sapien
5	1557.5	26.5	1170	1 ITAL HUMAN	P20701 homo sapien
6	1538.5	26.2	1163	1 ITAL MOUSE	P24063 mus musculus
7	1161.5	19.8	1167	1 ITAE MOUSE	Q60677 mus musculus
8	1153	19.6	1179	1 ITAE HUMAN	P38570 homo sapien
9	1103.5	18.8	1151	1 ITAL HUMAN	P56199 homo sapien
10	1093.5	18.6	1189	1 ITAH HUMAN	Q96155 homo sapien
11	1085	18.5	1170	1 ITAE BOVIN	P53710 bos taurus
12	1071	18.2	1178	1 ITA2 MOUSE	Q62469 mus musculus
13	1068	18.2	1181	1 ITA2 HUMAN	P17301 homo sapien
14	1064.5	18.1	1167	1 ITAG HUMAN	O75578 homo sapien
15	1062	18.1	1180	1 ITAL RAT	P18614 rattus norv
16	663	11.3	1039	1 ITAM MOUSE	Q00651 mus musculus
17	635	10.8	1038	1 ITAM HUMAN	P13612 homo sapien
18	630	10.7	1035	1 ITA9 HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA9 XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3 CRISP	P17852 cricetidae
21	567.5	9.7	1053	1 ITA3 MOUSE	Q62470 mus musculus
22	555.5	9.5	1053	1 ITA5 MOUSE	P11688 mus musculus
23	545.5	9.3	1050	1 ITA5 XENLA	Q06274 xenopus lae
24	542.5	9.2	1034	1 ITAV CHICK	P26008 gallus gall
25	537.5	9.1	1130	1 ITA6 HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITAV MOUSE	P43406 mus musculus
27	532.5	9.1	1066	1 ITA3 HUMAN	P26006 homo sapien
28	532	9.1	1049	1 ITA5 HUMAN	P08648 homo sapien
29	531.5	9.0	1072	1 ITA6 CHICK	P26007 gallus gall
30	526	9.0	1048	1 ITAV HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8 CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6 MOUSE	Q61739 mus musculus
33	498	8.5	1179	1 ITA7 MOUSE	Q61738 mus musculus

ALIGNMENTS

RESULT 1

ID	ITAM HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) {Leukocyte adhesion receptor Mol.}			
DE	{Neutrophil adherence receptor}.			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol. chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Plenning J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	496	8.4	1396	1	ITA2 DROME	P12080 drosophila
35	491.5	8.4	1146	1	ITAL DROME	Q24247 drosophila
36	491	8.4	1033	1	ITAB MOUSE	Q89000 mus musculus
37	489.5	8.3	1039	1	ITAB HUMAN	P08514 homo sapien
38	489	8.3	1025	1	ITAB HUMAN	P53708 homo sapien
39	486	8.3	126	1	ITAM CAVPO	P11578 cavia porce
40	476	8.1	1181	1	ITA7 HUMAN	Q13683 homo sapien
41	470	8.0	1106	1	ITA7 RAT	Q63258 rattus norv
42	467	7.9	1226	1	PAT2 CAEEL	P34446 caenorhabdi
43	445.5	7.5	1139	1	INAL CAEEL	Q03600 caenorhabdi
44	424	7.2	1115	1	ITA3 DROME	O44386 drosophila
45	385	6.6	1000	1	ITA5 DROME	Q9wlm8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
RT conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.M., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; W76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RMHULB.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGRC:6149; ITCAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; P:cell adhesion receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 4.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT

Query Match		99.5%;	Score 5846.5;	DB 1;	Length 1152;		
Best Local Similarity		99.2%;	Pred. No. 0;				
Matches 1128;		Conservative	7;	Mismatches	1;	Indels	1;
Qy	1	FNLDTENAMTFOENARGFGQSVVQLOGSRVVGVAPQEI	VAANQGRSLYQCYSTGSCBPI	60			
	17	FNLDTENAMTFOENARGFGQSVVQLOGSRVVGVAPQEI	VAANQGRSLYQCYSTGSCBPI	76			
Db	61	RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNT	VYKGLCFGLFGLNLRQOQOK	120			
	77	RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNT	VYKGLCFGLFGLNLRQOQOK	136			
Qy	121	FPFALRGCPQEDSDIAFLVDSGSIIPHDPRAKEPIST	WMBQLKSKTFLSLMOYSEEP	180			
	137	FPFALRGCPQEDSDIAFLVDSGSIIPHDPRAKEPIST	WMBQLKSKTFLSLMOYSEEP	196			
Qy	181	RIHFTFKFQNNPNSRLIKPTOLLGRTHATGIRKVVRE	LFINITNGARXNAFKILLI	240			
	197	RIHFTFKFQNNPNSRLIKPTOLLGRTHATGIRKVVRE	LFINITNGARXNAFKILLI	256			
Db	241	TDGEKFGPLGYEDVPEADREGVIRYVIGVDAPFRSE	KSRQELNTVASKPRDRHVFQIN	300			
	257	TDGEKFGPLGYEDVPEADREGVIRYVIGVDAPFRSE	KSRQELNTVASKPRDRHVFQIN	316			
Qy	301	NFEALKTIQNLREKIPALEGTOTGSSSPHEHMSQGE	FSAAITNGPILSTVGSYDAG	360			
	317	NFEALKTIQNLREKIPALEGTOTGSSSPHEHMSQGE	FSAAITNGPILSTVGSYDAG	376			
Qy	361	GVFLYTSKEKSTFINWTRVDSMDNDAVLGYAAAIL	ILNRVQSLVGLGAPRYQHIGLVANFR	420			
	377	GVFLYTSKEKSTFINWTRVDSMDNDAVLGYAAAIL	ILNRVQSLVGLGAPRYQHIGLVANFR	436			
Qy	421	QNTGWESNANVKTQIGAYPGASICSDVDVDSNGST	DLVLGAPHYEQTRGGQSVVCLP	480			
	437	QNTGWESNANVKTQIGAYPGASICSDVDVDSNGST	DLVLGAPHYEQTRGGQSVVCLP	496			
Qy	481	PGQARWQCDVAVLGEQCPNGRFGAALTVLGDVNG	DKLTDVALGAPCEEDNRCAGVILP	540			
	497	PGQARWQCDVAVLGEQCPNGRFGAALTVLGDVNG	DKLTDVALGAPCEEDNRCAGVILP	555			
Qy	541	HGTSGSISPSHSQRIAGSKLSPRIQYFGQSLGSG	GGDLTMOGLVBLTVGAQGHVLLLRSQ	600			
	556	HGTSGSISPSHSQRIAGSKLSPRIQYFGQSLGSG	GGDLTMOGLVBLTVGAQGHVLLLRSQ	615			
Qy	601	PVLRYKATMEFNPREVARNVFCNDQVVKGEKGE	SVRVLCHVQKSTRDLRGQIQSVVT	660			
	616	PVLRYKATMEFNPREVARNVFCNDQVVKGEKGE	SVRVLCHVQKSTRDLRGQIQSVVT	675			
Qy	661	YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQCT	ETLKLQLPNCIEDPVSPIVRLNF	720			
	676	YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQCT	ETLKLQLPNCIEDPVSPIVRLNF	735			
Qy	721	SLVGTPLSAFGNLRPVLAEDAQRLFTALPFFKNC	GNNDNICDDLSITPFSMSLDCLVVG	780			
	736	SLVGTPLSAFGNLRPVLAEDAQRLFTALPFFKNC	GNNDNICDDLSITPFSMSLDCLVVG	795			
Qy	781	GRFENVTVVRNDGDSVRTQVTPFFPLDLSVRK	VSTLQNGORSQSWELACESASSTEV	840			
	796	GRFENVTVVRNDGDSVRTQVTPFFPLDLSVRK	VSTLQNGORSQSWELACESASSTEV	855			
Qy	841	SGALKSTCSINHPIPPENSEVTFNITFDVDSKAS	LGNKLLKANVTSENMPRTNKEP	900			
	856	SGALKSTCSINHPIPPENSEVTFNITFDVDSKAS	LGNKLLKANVTSENMPRTNKEP	915			
Qy	901	QLELPVKYAVMVVTSRGVSTKYLNTASENTSRV	WQHGYQVSNLQGRSLPISLVLVVPV	960			
	916	QLELPVKYAVMVVTSRGVSTKYLNTASENTSRV	WQHGYQVSNLQGRSLPISLVLVVPV	975			
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVNVNCSIAVCQRIQCDIP	1020			
	976	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVNVNCSIAVCQRIQCDIP	1035			

RESULT 2
ITAM_MOUSE
ID ITAM_MOUSE STANDARD; PRT; 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN ITGAM.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CN NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
PX MEDLINE=88312584; PubMed=3044779;
RA Pycela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378(1988).
[2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
[3]
RN SEQUENCE OF 17-28.
RP MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07640; CAA30479.1; --
EMBL; M14293; AAR3494.1; --
PIR; S00551; S00551.
DR HSSP; P11215; IABX.
DR MGI; MGI:96607; Itgam.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; WFP_A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; WVFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; WVA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; WVEA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Calcium; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 1153 INTEGRIN ALPHA-M.
FT DOMAIN 17 1105 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1106 1129 POTENTIAL.
FT DOMAIN 1130 1153 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 164 350 WVEA.
FT REPEAT 337 400 FG-GAP 3.
FT REPEAT 401 452 FG-GAP 4.
FT REPEAT 454 515 FG-GAP 5.
FT REPEAT 517 575 FG-GAP 6.
FT REPEAT 580 632 FG-GAP 7.
FT CA_BIND 465 473 POTENTIAL.
FT CA_BIND 529 537 POTENTIAL.
FT SITE 1132 1136 GPPKR MOTIF.
FT DISULFID 66 73 BY SIMILARITY.
FT DISULFID 105 123 BY SIMILARITY.
FT DISULFID 654 711 BY SIMILARITY.
FT DISULFID 770 1023 BY SIMILARITY.
FT DISULFID 1028 1033 BY SIMILARITY.
FT CARBOHYD 58 58 N-LINKED (GLCNAC. .)
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .)
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .)
FT CARBOHYD 696 696 N-LINKED (GLCNAC. .)
FT CARBOHYD 734 734 N-LINKED (GLCNAC. .)
FT CARBOHYD 772 772 N-LINKED (GLCNAC. .)
FT CARBOHYD 801 801 N-LINKED (GLCNAC. .)
FT CARBOHYD 881 881 N-LINKED (GLCNAC. .)
FT CARBOHYD 907 907 N-LINKED (GLCNAC. .)
FT CARBOHYD 941 941 N-LINKED (GLCNAC. .)
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .)
FT CARBOHYD 994 994 N-LINKED (GLCNAC. .)
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. .)
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .)
FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .)
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. .)
SQ SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;
Query Match 76.1%; Score 4470; DB 1; Length 1153;
Best Local Similarity 73.9%; Pred. No. 5e-291;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;
QY 1 FNLDTEHMTQENARGFGQSVVQLQGSRRVVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTEHMTQENARGFGQSVVQLQGSRRVVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNNMGLSLAATTPPOLACGPTVHOTCSENYVKGCLPFLFGSNLRQOPQK 120
DB 77 PLQVPPAVNNMGLSLAVSTVPQOLACGPTVHONCKENTYVNGLCVLFGLNLRPQQ 136
QY 121 PPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEPTSTWEOQLKSKTLFSLMOYSEEP 180
DB 137 PPEALRECPQESDIVEFLIDGSGSINNIDPQKKEFYSTWEOQFKSKTLFSLMOYSEF 196
QY 181 RIHFTFEFONNPNRSLIKPIFOLLGRTHATGIRKVVRELFNITNGARKNAKILILI 240
DB 197 RIHFTFNDFKRNPSRSHVSPIKOLNERTKTSASIRKVVRELFHKTNGARENAKILVVI 256
QY 241 TDGEKFGDPLGYEDVIPADREGVIRVYVGDAPFRSEKSRQELNVTASKPRPRHVPQIN 300
DB 257 TDGEKFGDPLDYKDVIPADRAGVIRVYVGNAFNPKPSRRELDITASKPAGHEVFOVD 316
QY 301 NPEALKTIQONOLREKIPALEGTOTGSSSSPEHEMSQEGFSAAITSNGLLSTVGSYDNAG 360
DB 317 NPEALNTIQONOLQEKIPALBGTOTGSSSSPEHEMSQEGFSASITSNGLLSTVGSYDNAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNVSQSLVLGAPRYOHIGLVANFR 420
DB 377 GAFLYTSKDKVTFINMTRVDSMDNDAYLGASAVILNRNVSQSLVLGAPRYOHIGLVANFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 480
DB 437 ENPGTWEPHTSIKGSQIGSYFGASLCSVDMADGNTNULIGAPHYYEXTRGQSVSCPL 496
QY 481 PRGQSRWCDAVLYGEOQWGRGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAAYLF 540
DB 497 PRG-RARWQCEALLHGDQHPGRGAALTVLGDVNGDKLTDVAIGAPGEQENGAAYIF 555
QY 541 HGTSGGSIQSPHSORIIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 556 YGASTASLSASHSHRIIGAHFSPGLQYFGOSLGGQDLTMDGLMDLAVAQGHVLLRSQ 615
QY 601 PVLRKAIWEEFNPREVARNVTECNQVVKGEAGEVRVCLVHVKQSTRDLREGQIQSVVT 660
DB 616 PVLRLEATWEPSPKVARVACQOVLKNDAGEVRVCLRVKNTKDLREGDIQSTVT 675
QY 661 YDLALDSGRPHSRVNFETKSTRTOVLGTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB 676 YDLALDPVRSIRAFEDTKNTRERTQVGLMQKCECLKLILPCVDVDSVPIILRLNY 735
QY 721 SLVGTPLSAGNLRPVLAEADAORLFTALPPFPKNCNDNI CODDLSITPSFMSLCLVVG 780
DB 736 TLVGBPLRSFGNLRPVLANDAO RFTTAMPPFPKNCNDN SICODDLSITMSANGLDLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTPFFPLDLISYRKVSTLQONORSORSWL-ACESASSTE 839
DB 796 GPQDFMSVTLRNDGEDSYGTQVTVVYPSGLSYRSDASQNP LTKKPMFVKPFAESSSSSE 855
QY 840 VSGALKSTSCSINHPTFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTE 899
DB 856 GHGALKSTTWNINHPFPANSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTE 915
QY 900 FQLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHOYQVSNLQSRSLPISLVLP 959
DB 916 FQLELPVKYAVVMVTSDESSIRYLNTASEMTSKVIOHQYQFNILGQSRSLPVSVFWIP 975
QY 960 VRLNQTVIWDROPVTFSENLSSCTHKEPLPSHSDPLAELRKAPVNCVSIACORQCDI 1019
DB 976 VQINNVTVWDHPQVTFPSQNLSSACTEQKSPHSPRDLERTPVLNCVAVVCKRLQCDL 1035
QY 1020 PFFGIQEBFNATLKGNLSPDWYIKTSHNHLIVSTAELFNDSVFTLLPQCGAFVRSQTE 1079
DB 1036 PSFNTQEIFNVTLKGNLSPDWYIKTSHGHLIVSSTEILFNDSAPALLPQESYVRSKTE 1095
QY 1080 TKVEPEVNPPLIVGSSVGGILLIALITAAYLKLGPEKRYKQNMKSGGPGASPO 1137
DB 1096 TKVEPEVNPPLIVGSSIGGLVLLALITAGLYKLGFPRQYKQNMNEAAPQADAPPO 1153

RESULT 3
ITAX HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11C.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8816645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
a leukocyte adhesion glycoprotein, p150,95";
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
p150,95 molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M81695; AA55180.1; -.
CC EMBL; Y00093; CA68283.1; -.
CC EMBL; M29165; -; NOT ANNOTATED CDS.
CC EMBL; M29487; AA51620.1; ALT SEQ.
CC EMBL; M29482; AA51620.1; JOINED.
CC EMBL; M29483; AA51620.1; JOINED.
CC EMBL; M29484; AA51620.1; JOINED.
CC EMBL; M29485; AA51620.1; JOINED.

EMBL; M29486; AA51620.1; JOINED.
PIR; A36584; RWHUIC.
PDB; 1N3Y; 18-FEB-03.
Genew; HGNC:6152; ITGAX.
MIM; 151510;
GO; GO:0008305; C:integrin complex; TAS.
GO; GO:0004895; F:cell adhesion receptor activity; TAS.
GO; GO:0004872; F:receptor activity; TAS.
GO; GO:0007155; P:cell adhesion; TAS.
GO; GO:0007397; P:histogenesis and organogenesis; TAS.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWFA.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00057; Integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWFA_1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163 INTEGRIN ALPHA-X.
FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128 POTENTIAL.
FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 165 351 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 POTENTIAL.
FT CA_BIND 530 538 POTENTIAL.
FT CA_BIND 593 601 POTENTIAL.
FT SITE 1131 1135 GEFER MOTIF.
FT DISULFID 69 76 BY SIMILARITY.
FT DISULFID 108 126 BY SIMILARITY.
FT DISULFID 655 712 BY SIMILARITY.
FT DISULFID 771 777 BY SIMILARITY.
FT DISULFID 848 863 BY SIMILARITY.
FT DISULFID 998 1022 BY SIMILARITY.
FT DISULFID 1027 1032 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490 G -> A (IN REF. 2).
FT CONFLICT 756 756 L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;
Query Match 59.0%; Score 3469; DB 1; Length 1163;
Best Local Similarity 61.0%; Pred. No 5.2e-224;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;
Qy 1 FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60
Db 20 FNLDTELTAFRVDSAGFGDSVVQYANSVVVGAPQKITAAQNTGGILYQGYSTGACEPI 79
Qy 61 RLOVPVEANVMSLGLSLAATTSPPQLACGPTVHOTCSENTYVKGCLFLPGSNLRQOPQK 120
Db 80 GLQVPPVEANVMSLGLSLASTTSBQLACGPTVHHCGRNMYTGLCFLLPPT--QUTQR 137
Qy 121 FPEALRCGPQEDSDIAFLVDGSGSIIPHDFRAKEFTISTVMEQLKSKSLFLSLQYSEEF 180

138 LPVSRQECPRQEQDIIVFLDGGSSISRRNFATMNFVRAVISOQFQRPSTQPSLMQESNKF 197
181 RIHTFKBQNNPNPSLKIPITOLLGRTHATGIRKVVRELFNTNGARKNAFKILILI 240
198 QHTFTFEERFRTNPLSLLASVHQLQGTFTATAIQNVVHRLFHASGARRDATKILIVI 257
241 TDGKFGDPLGYEDVPEADREGVIRYVGVGDAPFSEKSRQBLNTVASKPRDHFQJLN 300
258 TDGKKGDSLDYKDVIPMDAAGIIRIYAGVGLAFQNRNSWKEINDIASKPSQEHFKVE 317
301 NFPAKTIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAALTSNGPLSTVSGVDMWG 360
318 DFDALXDIQNLREKIFAIEGTOTGSSSFEHMSQEGFSAVPTDGPVLGAVSGFTWSG 377
361 GVFLYTSKESKSTINRVDSDMNDAYLGVAARAILRNVSQSLVIGAPRYOHIGLVAMER 420
378 GAFLYPNNSPPTINMSQVNDVRDLSGLSTELALWGVQSLVIGAPRYOHTGRAVIFT 437
421 QNTGMWESANVKGITQIGYFASGLCSVDVDSNGSDTLVLIGAPHYYEQTRGQVSVCP 480
438 QVSRQWRMAEVTGTQIGYFASGLCSVDVDTGSDTLVLIGAPHYYEQTRGQVSVCP 497
481 PRQARWCDAYLYGEGQWGRFGAALTVDGVNGDKLTDVAIGARCEENRGAVYLF 540
498 PRQWR-RWKCWDAYLYGEGQWGRFGAALTVDGVNGDKLTDVIGAPGEBENRGAVYLF 556
541 HGTSGGISPSHSORLAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGQVHLLRSQ 600
557 HGVLGSPISPSHSORLAGSKLSPLQYFGQSLGGQDLTODGLVDLAVGARGQVLLRTR 616
601 PVLURVKAEMFNPFRVARNVFCNDQVWKEAGEVRLVCHVOKSTRDLRREGQISQVVT 660
617 PVLWVGVSQGFPAEIPRFAFECEQVWSEQTLVQSNICLYIDKRSKNLGSRDQSSVT 676
661 YDLALDSGRPHSAVNETKNSRTOVGLTQTCETLKLQNPNCIEPVPSPVILRLNF 720
677 LDIALDPGRLSPRATQSTQKRSLSRVVLGLAKHCENFNLLPSCEDSVPTILRLNF 736
721 SLVGTPLSAPGNRPVLAEDAORLFTALPPFERNQGNMDCODLSITFSFMSLDCLVVG 780
737 TLVGKPLAPRNLPRMLAALQRYFTASLPFEXKNCADHICQDNLGISFSPFGLKSLVG 796
781 GPREFNVTVVRDGDSDSYTOVTFPFLDLSVRKYSTLQNRQSRWSKLACASSTEV 840
797 SNLEMAEVWVNDGSDSYTOVTFPFLDLSVRKYSTLQNRQSRWSKLACASSTEV 854
841 SGALKSTCSINHPITPENSEVFNITFDVDSKASLGNKLLKANTYSENMPRYNKTFF 900
855 SQGTWSTSCRIHLIFRGCAITFLATEDVSPRAVLGDRLLLTANVSSENTPRTSKTF 914
901 QLELPKYAVVWVTSKYSTKYNLTAS-ENTSRVWQHOYOVNLSGQSLPLSLVLYP 959
915 QLELPKYAVVTVSSHQYTKYNLSESEKSHVAMGRYQVNNLQQRDLPLVSNFWVP 974
960 VRLNQTVINDRPQVTSSENLSTCTYKERLPSHDSFLAELRKAPVWNCVSIACQRIQCDI 1019
975 VELNQAQVWVNDGSDSYTOVTFPFLDLSVRKYSTLQNRQSRWSKLACASSTEV 1034
1020 PFGIOBEFNATIKHLSRPNYKTSNHLIIVSTAEILPNDVFTLLPQCAQVRSOTE 1079
1035 PPSVQBEELFTLKMFLSGVWVQVQILQKRVVSVVAEITPDTVSQVLPQCAQVRSOTE 1094
1080 TKVEPFPVNPPLIVGSSVGGLLALLALITAAALYKLGFFKRYQKDWSE 1128
1095 TVLEKYKVNPTPLIVGSSVGGLLALLALITAAALYKLGFFKRYQKDWSE 1143

DT 28-PEB-2003 (Rel. 41, last annotation update)
Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
ITGAD.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]_TaxID=9606;
[1] SEQUENCE FROM N.A.
TISSUE=Splice;
MEDLINE=96111956; PubMed=8777714;
Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
Stanton D.E., Gallatin W.M.;
"A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3";
Immunity 3:683-690 (1995).
[2]
SEQUENCE OF 1-235 FROM N.A.
MEDLINE=20187620; PubMed=10722744;
Noti J.D., Johnson A.K., Dillon J.D.;
"Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of S_{pl} and Sp₃";
J. Biol. Chem. 275:8959-8969 (2000).
[3]
SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
MEDLINE=96257236; PubMed=8666289;
Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
"Cloning and chromosomal localization of a novel gene encoding a human beta 2-integrin alpha subunit";
Gene 171:291-294 (1996).
[4]
INTERACTION WITH VCAM1.
MEDLINE=99059842; PubMed=9841932;
Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W., Hoffman P.A., Stanton D.E., Bochner B.S.;
"alpha₂ integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)";
J. Exp. Med. 188:2187-2191 (1998).
[5]
INTERACTION WITH VCAM1.
MEDLINE=99370002; PubMed=10438935;
Grayson M.H., Bochner B.S., Gallatin W.M., Stanton D.E.;
"The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1";
J. Immunol. 163:1984-1990 (1999).
-!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES FROM THE BLOOD.
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D ASSOCIATES WITH BETA-2.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A WFMA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE. FAMILY.
-!- SIMILARITY: Belongs to the integrin alpha chain family.
-!- SIMILARITY: Contains 1 WFMA domain.
-!- SIMILARITY: Contains 7 FG-CAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 954 Q -> QCVHGVEMQTSKQILCRPAGDAEHVQSGELPC
      PWCSEARFNIRAGPCR (in isoform 2).
      /FTID=VSP 902738.
      R -> W (IN REF. 1 AND 2).
      Y -> I (IN REF. 2).

FT CONFLICT 214
FT CONFLICT 660
FT STRAND 155
FT STRAND 164
FT TURN 165
FT HELIX 169
FT TURN 186
FT STRAND 191
FT STRAND 202
FT HELIX 208
FT HELIX 217
FT TURN 222
FT STRAND 229
FT HELIX 233
FT TURN 244
FT HELIX 247
FT TURN 250
FT TURN 253
FT STRAND 256
FT HELIX 274
FT TURN 277
FT STRAND 280
FT HELIX 288
FT HELIX 293
FT TURN 298
FT HELIX 300
FT HELIX 307
FT STRAND 311
FT TURN 318
FT HELIX 319
FT TURN 329
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.5%; Score 1557.5; DB 1; Length 1170;
Best Local Similarity 34.7%; Pred. No. 4.9e-96;
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

QY 1 FNLDPTENAMTFO--ENARGFGQSVVQLQSGSRVVVGAQPSIVAAANQSGLYQCDYSTGSC 58
DB 26 YNLDVVGARSFSPRAGRHFGYVLQV--GNGVIVGAPGE---GNSGSLYQCSQSTGHCL 81
QY 59 PIRLVQPVVEANMISGLSLAATSPQILLACGPTVHQTCSNTYVKGICFLFGSNLR--- 115
DB 82 PVTLR--GSNYTSKYLGMILTADPTDGSILACDGLSRTCDQNTYLSGLCYLFRQNLQGP 140
QY 116 -QQQKPFPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKGFISTVMEQLKSKTFLSLM 174
DB 141 LQGRGFCQECIKG-----NVDLVFLFDGSMISQDPDEFQKILDFMDKVMKLSNTSYQPA 196
QY 175 QYSEBFRHFTKFEQNNPNRSLIKPITQLGRTHATGIRKVVVRELFINITNGARKNAP 234
DB 197 QPSTSYKTFDFSDYVKRQDPALLKHVKHMLLTNTFGAINVATVFEVBELGARPDPAT 256
QY 235 KILILITGEKFGDPLGYEDVTPEDREGVIRYVLGVGDAPRSEKSRQELMTVASKPRD 294
DB 257 KVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGHFQTKESQETLHKFASKPASE 309
QY 295 FVYQINNFEALXKIQNLREKIFAIEGTQGTGSSSSSEHMSQEGFSNAITSNGPLLSTVG 354

```

```

DB 310 FVKILDTFEKLDLFTLQKIYVIEGTSKODLTSFNMELSSGSIADLGRHVAVGAVG 369
QY 355 SYDWAGGVF--LYTSKSEKSTFNMTRVDSMDNDAYLGVA--IILRNEVQSLVIGAPRYQH 412
DB 370 AKDWAGGFLDKADLQDDTFIGNELTPEVRAGYLVYTVTWLPSRQKSTLLASGAPRYQH 429
QY 413 IGLVAMPR--QNTGMWESNANVGTGIGAFGASLCSVDVDSNGSTDLVIGAPHYEQT 470
DB 430 MGRVLLFQBPQGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLIGAPLFYGEQ 489
QY 471 RGGQSVYCPPLRGORARWQCDV--LVGEQGPWGRGAALTVLGDVNGDKLTDVAICAP 528
DB 490 RGRVFIY-----QRRQLGFEEVSELOQDPYPLGRFGEAITALTDINGDLVDVAVGAP 544
QY 529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDQDLMGLVDLTV 588
DB 545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEQTQVLSGIGQWFGRSIHGVKIDEGDLADVAV 601
QY 589 GAQGHVLLFSQPVLRVKATMBFNPREVARNFECNDQVV--KGKEAGEVAVVCLHVQKSTR 647
DB 602 GAEQMIIVLSRPVDMVMTLMSFSPABI PVHEVECSYSTSNKKEGVNITICQI--KSLY 660
QY 648 DRLREGQIQSVVYDLDLSDGRPHSRAVENETNSTRQTVGLGTOTCEFLKLOLPRCI 707
DB 661 PQF--QGRLVANLTVTLQDGHRTTRRGLFPGGRHELRNIAVT--TSMSCDTSFHFVFCV 718
QY 708 EDPVSPVILNPSL---VGTPLS--AFGN-----LRPLAEDAQRLFTALFPPEKNCN 757
DB 719 QDLISPINVSLNFSLMEEBETPRDQRAQKDIPIILRPSLHSETWEI-----PPEKNGE 773
QY 758 DNICQDDLSITFSFMSLDCLVVGPRFNVTVTRNDGDSYRTQVTFPPPLDLISYRKVS 817
DB 774 DKXCEANLVSFSPARSRALRLTAFASLSVELSLNLEEDAYVYQLOLHFFPGLSPKRV 833
QY 818 TLQNRQSRWRPLACES--ASSTEVSGALSKTSCSINHPIPEPENSEVTENITPDVDSKAS 875
DB 834 ML---KPHSQIPVSCBELPESRLLSRAL---SCNVSSPIFKAGHSVALQOMFNTLVNS 887
QY 876 LGNKLLKANVTSENN---MPTNTEFQLELPVKTVAVTVVTVTSHGVSTKLYNFTASEN 931
DB 888 WEDSVELHANVTNNEDSDILEDNASATTI---IPILYINILIQDQSDSLYSFTPKGP 944
QY 932 TSVVMQHQYQV---SNLQGRSLP--ISLVFLVPLRLNQTVDNRPOVTFSENLSSTCHTK- 986
DB 945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVGVPPPPSEGPITHQMSVQMEPPV--PCHYED 1002
QY 987 -ERLPSSHSD--FLAELKAPVWNCIAVCQRIQCDIPFGIQEENATLKGNSLFDWYIK 1043
DB 1003 LERLPDRAEPCLPGALPRCPV-----FRQELLYQVIGTLELVGEIE 1044
QY 1044 TSNHLLIVSTAILFNDVSFTLLPGQGAFAVRQGTETKVPFPFVFNPLPLIVGSSVGGLL 1103
DB 1045 AS-SMPSLCSLSISFNSKHFHLYGSNASL-AQVVMKVDVYVEKQMLYLYLSGIGGLL 1102
QY 1104 LLALITAAALKYKAGFKKQYKDMMSG--GPPGAP 1136
DB 1103 LLLIFIVLYKVGFFKRNKLEKKEAGGVENGIP 1136

RESULT 6
ITAL MOUSE STANDARD; PRT; 1163 AA.
ID ITAL MOUSE
AC P24063, 1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
DB alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] _SEQUENCE FROM N.A.
 RP MEDLINE=91268576; PubMed=2051027;
 RX Kaufmann Y., Tseng E., Springer T.A.;
 RA "Cloning of the murine lymphocyte function-associated molecule-1
 RT alpha-subunit and its expression in COS cells.";
 RL J. Immunol. 147:369-374(1991).
 RN [2]
 RP SEQUENCE OF 24-42.
 RX MEDLINE=85188276; PubMed=3897182;
 RA Springer T.A., Teplow D.B., Dreyer W.J.;
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 RT glycoproteins and unexpected relation to leukocyte interferon.";
 RL Nature 314:540-542(1985).
 CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA,
 CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
 CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
 CC LEUKOCYTE RECRUITMENT.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A WFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 WFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M60778; AAA39426.1; -.
 DR PIR; I56126; I56126.
 DR HSSP; P20701; ILFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; WFA_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; WFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; WFA_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; WFA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium;
 KW Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 1163 INTEGRIN ALPHA-L
 FT DOMAIN 24 1084 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1085 1108 POTENTIAL.
 FT DOMAIN 1109 1163 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 39 88 FG-GAP 1.
 FT REPEAT 7 7 FG-GAP 2.
 FT DOMAIN 148 334 WFA.
 FT REPEAT 2 2 FG-GAP 3.
 FT REPEAT 399 454 FG-GAP 4.
 FT REPEAT 455 514 FG-GAP 5.
 FT REPEAT 516 573 FG-GAP 6.
 FT REPEAT 576 628 FG-GAP 7.
 FT CA_BIND 466 474 POTENTIAL.

FT	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GFPR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	108	126	BY SIMILARITY.
FT	DISULFID	147	159	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1163	1163	N-LINKED (GLCNAC. .) (POTENTIAL).
SEQ	SEQUENCE	1163	AA; 128343	MW; A7A30784898232F CRC64;

Query Match 26.2%; Score 1538.5; DB 1; Length 1163;
 Best Local Similarity 34.2%; Pred. No. 9.1e-95;
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

QY	1	FNLTENAMTPQENA-RFGQSVVQLQGRVVCAPQBIIVAANQGRSLYQCDYSTGSCPE	59
DB	24	YNLDTRPTQSLAAGRHFGVQLQIEDG-VVVGAGE---GNTGGLYHCTSEFCQP	79
QY	60	IRLOPVEAVNMISLGLSLAATSPPLLACQPTVHQTCSENYYVGLCFLPGSNLRQOPQ	119
DB	80	VSLH-GSNHTSKYLGWTLATDAAGKSLACDPLGSLRTCDQNTYLSGLCYLPQSLGEM	138
QY	120	KFPALRGCPEDSDIAFLVDGSGSIIPHDPRAKEPISITVMEQLKSKTLPFSLMOYSEB	179
DB	139	QNRPAYQECMKGVLDLFLFDGSGSLDRKDFEKLFEKLVKDWVKLSINTSYQFAAVQST	198
QY	180	FRIHFTFKEF-QNNPNPRSLTKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILI	238
DB	199	CRTEFTLDYVVKQKNPDVLGSGVQPMFLINTFRALNYVVAHVFKESGARPDATKVLV	258
QY	239	LITDGEFGDPLGYEDVIPADREG-----VIRYVIGYGDAPRSKRSQELATVASKP	291
DB	259	IITDG-----EASDKGNISSAHDITRYIIGIKRHFVSVQKQKTLHFASEP	304
QY	292	PRDHVFOINPEALKTIQNLREKIFAIEGTQTCSSSFEHEMSQEGFSAITNGPILLS	351
DB	305	VEEFVKILDTFEKLDLFTDLQRIYALEGTRQDLTSFNELSSSGISADLSKGAHVVG	364
QY	352	TVGSDYDAGGVF-LYTSKESKSTFNTTRVDSMDNDALGYAAA-IILNRVQSLVIGNPR	409
DB	365	AVGAKDAGGFLDLREDLQCATFVQBPGLTSDVGGYLVTVAMWGRSSRSLPLAAGAPR	424
QY	410	YOHIGLVAMPR-QNTGWMESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLICAPHY	467
DB	425	YOHVGVLLFOAPEAGGRWNQTKIEGTQIGSYFGSELCSVDLDQDGEALLIGAPLFF	484
QY	468	EOTRGQSVSCPLPRGQARQCDAVLYGEGQDPWGFAGALTVLGVDNGDKLTVGAIGA	527
DB	485	GEQRGRVFTY---QRRQSLFEMVSELQGDGPGYPLGFGAATALTALTDINGDLTVA	541
QY	528	PGEEDNRGAVYLPFGTSSSGISPSHSORISAGSLKSLPRLOYFGQSLSGGQDITMDGL	587
DB	542	PLEE-QGAVYIFNGKPG-GLSPQPSQRIQGAQFPGIRWFGRSIHGVDLGGRLADV	598
QY	588	VGAQGHVLLLSQPVLRYKAIMTFNPREVARNVPECDNQVYVKGKAG-EVRVCLHFVKST	646

Db 599 VAGRVRVLSRPVVDVVTLSFSPBEIPVHEVECSASAREBQKHGVKLKACRIFPLT 658
Qy 647 RDRREGQIQSVVYDLDALDGRPHSRAVFNENKTRRQVGLTQTCTBLKQLPNC 706
Db 659 PQ--PQGRLLANTSLYQLDGRNRSGRLFPGGHSLSGNTSITP-DKSLCDFHFPIC 715
Qy 707 LEDPSPVILNLSLV---GTPLSAFGN-LRPVLAEADAQLFALPFPKNGCNDNIQ 762
Db 716 IODLSPINVLNLSLEBEGTPRQKRAMQPIILRPSIHVV-TKEIPFKNCGEDKKCE 774
Qy 763 DDLSTFTFMSLDCLVWGP-----REFNTVTVTRNDGSDSYRTQVFPFPLDLSYRKV 816
Db 775 ANLTSSPARS-----GPLRLKSSASLAVENTLNSGSDAYVVRDLDPFRLGSPRKV 827
Qy 817 STLQNRQCRWRKLACASSTVSGAL-KSTCSINHPDPENSEVTFTNTFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCEBL--TEGSLSTKLKCNVSPFPKAOEVSLOQMENTLLNS 882
Qy 876 LGKLLKLLKANTSEN-NMPTNKTEFQLELPKAVYVWVTSVSHGVSTKYLNFTASENTR 934
Db 883 WEDFVNLNTHVCEENSLQEDNSAATHIPVLPVNLTKQENSTLYISFTPKGPKTQ 942
Qy 935 VQHOYQVSNLQGRSLPISLVPLVRLNQTVIDRPO-----VTFSENLS--TCHTK 986
Db 943 QVQHYVQV-----RIQPSAYDHNMT-LEALVGVPRPHSEDLITYVTSVQTDPLVTCHSE 996
Qy 987 E-RLPSSHDFLAELRKAIPVNCISIAVCORICDIPFGIOGBFNATLKNLSFDWYIKTS 1045
Db 997 DKRPSE---ABQCPLGV-----QRCPIVF---RWEILLQVGTVELSKELKAS 1042
Qy 1046 HNHLIVSTAEILFNDSTFTLPGGAPVRSQETKVPFVPPNPPLPLVSSVGGLLLL 1105
Db 1043 -STLSLCSLSVSFNSKHFLYLSKA-SEAQLVKVLDLHEKMLVYVLSGIGLVLL 1100
Qy 1106 ALITAAVLYKLFKQYKDMV-SEGPPCAEP 1136
Db 1101 FLIFLALYKVGPFKRLKEMADGGVNGSP 1132

RESULT 7
ITAE MOUSE
ID ITAE MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC STRAIN=AKR.
RX MEDLINE=95187992; PubMed=7982170;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Weis J.H.;
RT Immunity 1:393-403(1994).
RL -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: type 1 membrane protein.

-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
-!- SIMILARITY: Belongs to the integrin alpha chain family.
-!- SIMILARITY: Contains 1 VWFA domain.
-!- SIMILARITY: Contains 7 VWFA repeats.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U12236; AAC52142.1; -.
DR HSP; P11215; IABX.
DR MGD; NGI:1298377; Itgae.
DR InterPro; IPR000413; integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 3.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Magnesium;
Calcium.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 1167 INTEGRIN ALPHA-E.
FT CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1115 1137 POTENTIAL.
FT DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 193 384 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 449 501 FG-GAP 4.
FT REPEAT 503 564 FG-GAP 5.
FT REPEAT 566 631 FG-GAP 6.
FT REPEAT 634 686 FG-GAP 7.
FT CA_BIND 514 522 POTENTIAL.
FT CA_BIND 578 586 POTENTIAL.
FT CA_BIND 646 654 POTENTIAL.
FT DOMAIN 185 191 GLU-RICH (ACIDIC).
FT SITE 1140 1144 GFGR MOTIF.
FT DISULFID 72 83 BY SIMILARITY.
FT DISULFID 130 164 BY SIMILARITY.
FT DISULFID 698 754 BY SIMILARITY.
FT DISULFID 814 820 BY SIMILARITY.
FT DISULFID 884 898 BY SIMILARITY.
FT DISULFID 998 1023 BY SIMILARITY.
FT DISULFID 1031 1047 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .)
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .)
FT CARBOHYD 314 314 N-LINKED (GLCNAC. .)
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .)
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .)
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .)
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .)
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .)
FT CARBOHYD 773 773 N-LINKED (GLCNAC. .)
FT CARBOHYD 829 829 N-LINKED (GLCNAC. .)
FT CARBOHYD 846 846 N-LINKED (GLCNAC. .)
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .)
FT CARBOHYD 925 925 N-LINKED (GLCNAC. .)
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .)

FT CARBOHYD 1013 1013 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 1055 1055 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 1086 1086 N-LINKED (GLNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCFFD CRC64;

Query Match 19.8%; Score 1161.5; DB 1; Length 1167;
 Best Local Similarity 29.0%; Pred. No 1 6e-69;
 Matches 357; Conservative 214; Mismatches 457; Indels 205; Gaps 43;

QY 1 FNDLTENA--MTQENARFGQSVVQLOGRVVGAPQIIVANORGS-----LYQCDY 52
 DB 20 FNDVDMWVWYALQPCAPVLSLLHDPEN-----NOTCLLVARESSNNTAALYRCAL 74
 QY 53 STGSCPIRLQVPEAVNNSLGLSLAATT--SPQLLAC-GPTVHQCSTENTYVKGCLFL 109
 DB 75 SI-SPDEIACQ-PVEHICMPKRGYQGVTLVGNHNGVLCVQVQARFNSINSELTGACSL 132
 QY 110 FGSNLRQQQQKPFPEALRG-----C-----POE 131
 DB 133 LTPNLDLOAQAVFSDLEGFLDPCAHVDSGDYCRSGKSGTECKSARRRRTVEEDED 192
 QY 132 DSIAFLVDGSGIIPHDPRAKEFTSTVMEQL--KSKTFLSLMOYSEEPRIHFTFEF 189
 DB 193 GTBIAIVLDGSGIGSDFOKAKNFITMTNRYEKCFCNFAVQYGAIVOTEPDLOES 252
 QY 190 QNNPNRSLIKPTQLGRTHRTATGIRKVVRELFNITNGARKNAFKILILITDGEKFGDP 249
 DB 253 RDINASLAKVQSVQVKEVTKTASAMQVLDNIFIPSRGSRKALKVMVLTGDIKFGDP 312
 QY 250 LGYEDVPEADRGVRYVLCVGDARFSEKSRRELATVASKPRDHFVQINNEPEAKTIQ 309
 DB 313 LNLTTVINSKMGQVRFVFAIGVGRFKNNTYRELKLIASDPRKXHTFFKVTNYSALDGLL 372
 QY 310 NQLEKIFALGEGTQSSSFHEMSQBSAITSNGP--LLSTVSGSYWAGVFLY-TS 367
 DB 373 SKLQORVHMEGT---VGDALQVLAQTGFSAQILDKGVLLCTVGFANWGGGALLYSTQ 429
 QY 368 KEKSTFINMT-RVDS-DMNDAYLGIAAAIILNRVQSLVGLGARYQHIGLVAMFRONTGM 425
 DB 430 NGRGCFINOTAKEDSRVQVSYLGSLAVLKHAGISYVAGAPRRKLRGAVFLERKEDR- 488
 QY 426 WESNA---NVKGTGIGAYFCASLCSDVDNSNGSTDLVLICAPHYETRGQGVSVCLPR 482
 DB 489 -BEDAFVRIEKGQMGYFVSCLPVDIDMDGTTDFLLVAPPHYHIRGEGRVYVQVPE 547
 QY 483 GQARWQCDVAVGEQOQPMGRFGAALTVLGDVNGDKLTDAVIGAP-----GRENPGA 536
 DB 548 -QDASPSLAHTLSGHPGLTNSRPGFAMAAGVINDQKFTDVAIGAPLEGFGAGDGASYGS 606
 QY 537 VYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGSLSGGDLTMDGLVLTGVAQGHVLL 596
 DB 607 VYIYNGHSG-GLYDPSQQTIRASSVAGSLHVFMSVSGGLDFNGDGLADITVSRDSAVV 655
 QY 597 LRSQVPLRVKAIMBPNPREVARNVFECDNQVVKGEAGYRVCLHVOKS---TRDELREG 653
 DB 666 LRSRPVVDLTVSMFTFP-----DALPMVFIKGM--DVNLCPVDVSVVASEPGLREM 715
 QY 654 QIQSVVYDLDALDSGRPHSRVAFNETKSTRTQVTLGLTQC-----696
 DB 716 FLNFTVDV-----TKQRQLQCESGCGQSLRKNWGSFICEHPWLI 760
 QY 697 ETKLQLPNCIEDVPSVILRLNFSVLGTPLSAPGNLR---PVLARDQALFTALP--P 750
 DB 761 STEEL-----CEDCESNITIKVYE-----PQTSGRRDYPNPL--DHYKEPSAIFQLP 809
 QY 751 FEKVCNDNIQDDLSITFFSMGLDCLVWGSPRENTVTVRNDGESYRTQTFFPPLD 810
 DB 810 YEKDCKRVFCIABIQTLTN--ISQQLVWGVTKVETWNISLTNSGDSYMTNMAIYPRN 868
 QY 811 LSVKSVTLQNRQSRNRLACBSASSTEVSGALKSTSCSINHPIFFENSEVPIFNTEFV 870
 DB 869 LQFKKI-----QKPSVDPVQDDPKPV---ASVLVWNCIKIGHFIL-KRSVNVSVTWQL 918

QY 871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNTFASE 930
 DB 919 BESVFPNRTADITVTISNSNEKSLARETR---SLQFRAFIATVLSR--PSVMYWN--TSQ 971
 QY 931 NTSVMQHQYQVSNLQGRSLFISLVFLVPLNQTVMWDRPQVTFSENLST-----CHT 985
 DB 972 SPDSKKEPFFNVHGENLFGAVFQIQICVPIKQDF-----QIVRVKMLTKTQDHTECTQ 1025
 QY 986 KERUPSHSDFLAELRKAPVNCSTAVCQRIQCDIPFGIQEENATLKNLSFDWYIKTS 1045
 DB 1026 SOEPACGSDPQVQKWHSHVVCAL-----TSNKENTVVAEISVVG 1065
 QY 1046 RNHLILVSTA-----EILFNDVSFTLLPGQAFVRSQSTETKVPFF-----EVPNPLPIV 1095
 DB 1066 HTKQLLRDVSBLPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEBETRSLPLII 1119
 QY 1096 GSSVGLHLLALITAAALKYKLGFKKQYKDMKSE 1128
 DB 1120 GSSIGGLLVVITAILFKCGFKKYOQLNLE 1152

RESULT 8
 ITAE HUMAN
 ID ITAE HUMAN STANDARD; PRT: 1179 AA.
 AC P8570; Q9NZU9;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
 DE antigen) (CD103 antigen) (Integrin alpha-IEL).
 GN ITGA8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
 RC TISSUE=Leukemia, and Lymphocytes;
 RX MEDLINE=94164962; PubMed=8119947;
 RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
 RA Parker C.M.;
 RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E
 RT subunit. Unusual structure and restricted RNA distribution.";
 RL J. Biol. Chem. 269:6016-6025(1994).
 RN [2]
 RP REVISIONS TO 88-114.
 RA Parker C.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 53-1179 FROM N.A.
 RC TISSUE=Petal kidney;
 RX MEDLINE=20138496; PubMed=10673275;
 RA Touchman J.W., Ankster Y., Dietrich N.L., Maduro V.V., McDowell G.,
 RA Shotelsius V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
 RA Green E.D.;
 RT "The genomic region encompassing the nephropathic cystinosis gene
 RT (CTNS): complete sequencing of a 200-kb segment and discovery of a
 RT novel gene within the common cystinosis-causing deletion.";
 RL Genome Res. 10:165-173(2000).
 RN [4]
 RP MUTAGENESIS OF ASP-109 AND PHE-316.
 RX MEDLINE=20400502; PubMed=10837471;
 RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
 RA Brenner M.B.;
 RT "The role of alpha and beta chains in ligand recognition by beta 7
 RT integrins";
 RL J. Biol. Chem. 275:25652-25664(2000).
 CC -I- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
 CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
 CC CELL MONOLAYERS.
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
 CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L25851; AAB59359.2; -;
 CC EMBL; AF168787; AAF43107.1; -;
 CC PIR; A53213; A53213.
 CC HSSP; P11215; IABX.
 CC Genew; HGNC:6147; ITGAE.
 CC MIM; 604682; -;
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWFA.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 3.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS02343; WFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Signal; Repeat; Polymorphism; Magnesium; Calcium.

FT CHAIN 19 1179 INTEGRIN ALPHA-E.
 FT CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.
 FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
 FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
 FT TRANSFEM 1125 1147 POTENTIAL.
 FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 181 198 GLU-RICH (ACIDIC).
 FT REPEAT ? ? FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).
 FT DOMAIN 200 391 WFA.
 FT REPEAT 401 456 FG-GAP 3.
 FT REPEAT 457 506 FG-GAP 4.
 FT REPEAT 510 571 FG-GAP 5.
 FT REPEAT 573 638 FG-GAP 6.
 FT REPEAT 641 693 FG-GAP 7.
 FT CA_BIND 522 530 POTENTIAL.
 FT CA_BIND 586 594 POTENTIAL.
 FT CA_BIND 654 662 POTENTIAL.
 FT SITE 1150 1154 GPPR MOTIF.
 FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 126 159 BY SIMILARITY.
 FT DISULFID 706 762 BY SIMILARITY.
 FT DISULFID 823 829 BY SIMILARITY.
 FT DISULFID 893 907 BY SIMILARITY.
 FT DISULFID 1008 1033 BY SIMILARITY.
 FT DISULFID 1041 1057 BY SIMILARITY.
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).

726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
 360 360 D -> E.
 1041 1041 /FTID=VAR_008884.
 208 208 C -> S.
 316 316 /FTID=VAR_008885.
 477 477 D -> A: LOSS OF E-CADHERIN BINDING.
 482 482 V -> I (IN REF. 3).
 482 482 Q -> R (IN REF. 3).
 950 950 R -> W (IN REF. 3).
 1019 1019 R -> W (IN REF. 3).
 1179 1179 A -> V (IN REF. 3).
 AA; 130088 MW; E558902ED9D95E1 CRC64;
 Query Match 19.64; Score 1153; DB 1; Length 1179;
 Best Local Similarity 29.14; Pred. No. 5.9e-69;
 Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;
 45 GSYQCYSTGS--CEPI-RLQVP-----VEAVNMSGLSLAATTSPQLACGPTVHQ 95
 65 GPLHRCSLVQDEILCHPVEHVPDPKGRHGVTVVRSRKGVLICI-----QVLVRP--HS 117
 96 TCSENTYKGLCFEGSNLRQPO-----SDIAFLVPGSGSIIPDFRAKEPISTVMEQL-- 164
 118 LSSELT---GTCSLGLPDLRPOAQANFDLENLDPDARVDTGDCYSNKEGGEDVNTA 174
 120 KPFEALRGCPQED-----SIAFLVPGSGSIIPDFRAKEPISTVMEQL-- 164
 175 RQRALKEEEDKEEEDDEEAEAGTEIAILDGGSIDPDPFORAKDFISNMNRNFE 234
 165 KSKTILFSLMOYSEPRHFPTFKFQNNPRLSKIPITOLLGRTHATGIRKVVRELFN 224
 235 KCFECNFALVQGVQIQTPEFLRSDQVNASLARVQNTQGVSVTKTASAMOHVLSIPT 294
 225 ITNGARKNAFKILILITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSR 284
 295 SSHGSRKASKWVVLTDGGIFEDPLNLTVINSPQGVKVERFAIGVGBEBSKARTARE 354
 285 NTVASKPRDHVQINNEALKTIQNLREKIFALEGTQGTSSSSFEHMSQEGSAAT 344
 355 NLIASDPDETHAFKVTNMDGLSKLRYNIISMEGT---VEDALHYLAQIGFSAQL 411
 345 SNGP--LLSTVGSYDWAGGVLY--TSKSKSTFINTRVDSMDMDA----VLGYAAAILRN 398
 412 DERQVLLGAVGAPDWSGALLYDTSRRGRFLNQTAADAADEAAQYSLGVAVAVLHKT 471
 399 RVQSLVLGAPRYQHIGLVAMPR--QNTGWESNANV--KQTQICAYFGASLCSVDVNSG 456
 472 CSLSVVAGAPQYKHG--AVPELOKEGREASPLVLEGEQMGSYFGSELCPVIDMDG 529
 457 DLVLIGAPHYEOTRGQGVSCFLPRGORARWOCDAVLYGEOGQWGRPGAALTVLGDVN 516
 530 DFLVAAPFYHVGEGEGRVYVTRLSE--QDGSFSLARILSGHGFNTNARFGFMAAME 588
 517 GPKLTDVAIGAP--GEEDNR--GAVYLFHOTSGSGISPSHSQIRIAGSKLSPRLOYP 570
 589 QDKLTDVAIGAPLEGEGDADGASFGSVIYING--HWDGLSASPSQIRIASTVAPGL 647
 571 SLUSGGDLTMOGLVLTGCAQGHVLLRSQVLRKYKAIMENPRVARNVPCNDQVVK 630
 648 SMAGGFDSGDGLADITVGTLCQAVVFRSRPVVRUKVSNAPTTPSALP-----IC 697
 631 KEAGEVRVCLRVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRVAVENKSTRQTV 690
 698 NGVNVRLCFEI--SSVTTASGESGLREALNFTLDVVGKRRRLQCSQVRSGLCGL 756
 691 GLTQTCETLKLQLPN-----CIEDPSPVIVRLNFSLVGTPLSAFGNLRVPLAEDA 746

	GO: GO:0005518; F.collegen binding; TAS.	
DR	GO: GO:0007160; P.cell-matrix adhesion; TAS.	
DR	GO: GO:0007517; P.muscle development; TAS.	
DR	InterPro: IPR000413; Integrin_alpha.	
DR	InterPro: IPR002035; WVF A.	
DR	Pfam: PF01839; FG-GAP; 3_	
DR	Pfam: PF00092; vwa; 1.	
DR	PRINTS: PR01185; INTEGRINA.	
DR	PRINTS: PR00453; VWFADOMAIN.	
DR	SMART: SMO0391; Intc_alpha; 5.	
DR	SMART: SMO0327; VWFA; 1.	
DR	PROSITE: PS00242; INTEGRIN_ALPHA; FALSE_NEG.	
DR	PROSITE: PSS0234; WVFA; 1.	
KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;	
KW	Signal; Repeat; Calcium; Magnesium; Polymorphism.	
SIGNAL	1 22	POTENTIAL.
FT CHAIN	23 1189	INTEGRIN ALPHA-11.
FT DOMAIN	23 1142	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	1143 1165	POTENTIAL.
FT DOMAIN	1166 1189	CYTOPLASMIC (POTENTIAL).
FT REPEAT	38 94	FG-GAP 1.
FT REPEAT	102 163	FG-GAP 2.
FT DOMAIN	167 345	VWFA.
FT REPEAT	359 420	FG-GAP 3.
FT REPEAT	422 475	FG-GAP 4.
FT REPEAT	477 537	FG-GAP 5.
FT REPEAT	539 598	FG-GAP 6.
FT REPEAT	601 653	FG-GAP 7.
DOMAIN	1154 1162	POLY-LBU.
DOMAIN	1174 1177	POLY-ARG.
CA BIND	488 496	POTENTIAL.
CA BIND	551 559	POTENTIAL.
CA BIND	613 621	POTENTIAL.
DLSULFID	76 83	BY SIMILARITY.
DISULFID	121 139	POTENTIAL.
DISULFID	129 159	POTENTIAL.
DISULFID	659 668	BY SIMILARITY.
DISULFID	674 729	BY SIMILARITY.
DISULFID	781 787	BY SIMILARITY.
DISULFID	881 893	BY SIMILARITY.
CARBHYD	82 82	N-LINKED (GLCNAC..)
CARBHYD	95 95	N-LINKED (GLCNAC..)
CARBHYD	291 291	N-LINKED (GLCNAC..)
CARBHYD	331 331	N-LINKED (GLCNAC..)
CARBHYD	358 358	N-LINKED (GLCNAC..)
CARBHYD	449 449	N-LINKED (GLCNAC..)
CARBHYD	462 462	N-LINKED (GLCNAC..)
CARBHYD	528 528	N-LINKED (GLCNAC..)
CARBHYD	642 642	N-LINKED (GLCNAC..)
CARBHYD	694 694	N-LINKED (GLCNAC..)
CARBHYD	857 857	N-LINKED (GLCNAC..)
CARBHYD	894 894	N-LINKED (GLCNAC..)
CARBHYD	973 973	N-LINKED (GLCNAC..)
CARBHYD	1032 1032	N-LINKED (GLCNAC..)
CARBHYD	1040 1040	N-LINKED (GLCNAC..)
VARIANT	433 433	V-> M. /FTId=VAR_009889.
VARIANT	524 524	R -> L. /FTId=VAR_009890.
VARIANT	972 972	L -> P. /FTId=VAR_009891.
VARIANT	1003 1003	I -> M. /FTId=VAR_009892.
VARIANT	1030 1030	Missing. /FTId=VAR_009893.
VARIANT	1094 1094	L -> V. /FTId=VAR_009894.
SEQUENCE	1189 AA; 133609 MW; 60303C08AA4CD52 CRC64;	

Query Match 18.6%; Score 1093.5; DB 1; Length 1189;
Best Local Similarity 28.2%; Pred. No. 5.8e-65;
Matches 348; Conservative 215; Mismatches 502; Indels 167;

Qy	1	BNLDTENAMTFOENARG-FCQSUVQ---LQSR-VVVGAPQEIIVAANQBSGLYQCDYSTGS	56
Db	23	FNMIDTRKRPVIPSRTAFPGTYYQOHDISGNKMLVVGAPLETNGYQKTGDVYKCPVHGN	82
Qy	57	CEPRL-----QVPBEAVNMSLGLSLAAATTPPQILLACGPTVHQTCSNTYVYKGLCFLF	110
Db	83	CTKLNLRVTLNVSRKDNWRLGLSLATNPKNSPLACSPLWSHCSSYTTGMSRV	142
Qy	111	GSNLRQPOKPEALRCPOEDSDIAPLVQSGSIIPHDPRRKEFTSWEO--LKSK	168
Db	143	NSNFRFSKTWAP-ALQKC-QTYMDIVTLQSGNSIY-WEVQHPFLNLKCYIGPGQ	198
Qy	169	TLPSLMQYSEBFRHFTKEFQNNENPRSLIKPITQLLG-RTHTATGIRKVVRLBNITN	227
Db	199	IQGVVQYGEDVWEHFLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEPARSAFO--K	256
Qy	228	GARKNAPKILLITDGKFKGDPGLGYEDVPEADREGVIRVIGV-----GDAFRSEKSRQ	282
Db	257	GGRKAGKKNWIVITDGHSDP-DLEKVIQOSEDNVTREAVAVLGYNNRGINPEFLN	315
Qy	283	ELMTVASKPRDRHVQINNPEALKTIQNLREKIFAIEGTQSGSSSPEHMSQEGSAA	342
Db	316	EIKYIASDPDDKEFFNVITDEAALKDIVDALCDRIFLSEGTNK-NETSFGLMSQTGFSSH	374
Qy	343	ITSGPILLSVGSYDAGGVPLVTSKEK-----STFINMTRVDSMDNDAYLGYAAAIILR	397
Db	375	VVEDGVLLGAVGAYDNAGVILKETSAGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVVS	434
Qy	398	NRV-QSLVLGAPRYQHTGLVAMP-RONTGMWSSNANVKGTQIGAYFGASLCSVDVDSNGS	455
Db	435	SROGRVTVAGAPRPNHTGKVLFTMHNNRSUTTHQNMKGQOIGYFSGEITSVDIDGCV	494
Qy	456	TDVLIGAPHYEYTR--GGQVSVCPLPRGORARWQCDVILGYEQCPWGRFGAALTVLGD	514
Db	495	TDVLLVGAPMYFNEGREKGVVYEL---RQNRVYNGTLKDSHYNQARFGSSIASVRD	551
Qy	515	VNGDKLTVAIGAPGEDNRGAVLFTGTSQSGISPSHSQRIAGSKLSPLROYGQSLSG	574
Db	552	LNQDSYNDVVVGAPLEDNHAGAIYIFGPRGS-ILKTPKORITASELATGLQYFGCSITH	610
Qy	575	QDGLTMDGLDVLTVGAQCHVLLRSPQVLVKAIEMFNPREVARNVF--ECNQDVVKKE	632
Db	611	QLDLNEDGLIDLAVGALGNVILWSRFVQINASHFEPSKI--NIFHRDC-----KR	661
Qy	633	AGEVRVCL-----HVQSTRDLRLBEOIQSVVTYDIALDSGRPHSAVFNET	679
Db	662	SGRDATCLAAFLCPTPIFLAPHFQTTVG-----IRYNATMDERRVTPRAHLDEG	711
Qy	680	KNS--TRRQTOVLGHTQCTELKQLPNCIEDPVSPIVLRNLFSLVGTPLSAGMLRPVIA	738
Db	712	GDRFTNRAVLLSSQELCERINFHVL-DADYVKPVTFSVEYSLEDP-----DHGPMID	764
Qy	739	EDAQRLLTALPPFKKNCNDNICODDL-----SITFSF	771
Db	765	DGMPTTLRVSVFPNGCNEDEHCVDPDLVLDARSPLTAMEVCQVRILKPAQDCSAYTLSP	824
Qy	772	MSLDLVVGGPREFNVTVTVRNDGEDSYRTQVTFPFDLSYRKVSTLQNORSQSRWRLA	831
Db	825	DTTVPIIESTRQVAVENTLENRGANYSTVLNITSQSANLQF--ASLIQKEDSDGS--IE	880
Qy	832	CESASSTEVSGALKSTSCSINHPTFPENSEVTNITFDVDSKASLGNKLLKANVTSENN	891
Db	881	CYNBER-----RLQKQCNVSYPPFRKAKVAFRLDDEF-SKSLFLHLSIELAAGSDSN	934
Qy	892	MPRTNKTE--FQLBPLVKIAYVMVTVSHGVSTKY---LNFITAS--ENTSRVMQHVQSVN	944
Db	935	ERDSTKDNVAPLRFHLKYEADVLFTRSSLSLHVEVKLNSSLERYDGIQPPFCIFRION	994
Qy	945	LGQ---RSLPISLIVPLVPVLLAQTVWDRPQVTFSENLSTTC---HTKERLPSHSDFLA	997
Db	995	LGLPEFHGHMMKJITIPATISGNRLKLRLDELT-DEVANTSCNIGWSTETRYTPVE--B	1051
Qy	998	ELRKAPVYNSIAVCORIQCDIPFGIQEBEFNATIKGNLSDMY-----IKTSHNLLIV	1051

Db 1052 DRRAPQLKSHSDVVSINCNRLVP-NOEINFLHGLND---WLSRLKALKYKSKMIWN 1107
 Qy 1053 STAEILFNDSVFTLLPGQAFVRSQTKVEFEVFN-----PLPLVGVSGVGLLLA 1106
 Db 1108 AALQRFH-SPF-----IFREEDPSQIVFELSKQEDWQVPTIIVGSLGGLLLA 1158
 Qy 1107 LITAAIYKLFKPK-ROYKDMXSEGPPCAEP 1136
 Db 1159 LVLALWKLGFPRSAARRRE-----PGLDP 1183

RESULT 11
 ID ITA2 BOVIN STANDARD; PRT; 1170 AA.
 AC P53710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
 GN ITGA2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=7511592;
 RA Kamata T., Puzon W., Takada Y.;
 RT "Identification of putative ligand binding sites within I domain of
 RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663(1994).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC -!- ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L25886; AAB59255.1; -.
 CC PIR; I45914; I45914.
 CC HSP; P17301; LAOX.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWFA.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC SMART; SM00191; Int_alpha; 5.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Receptor; Polymorphism; Calcium; Magnesium.
 FT NON_TER 1 1

FT SIGNAL <1 18
 FT CHAIN 19 1170
 FT DOMAIN 19 1121
 FT TRANSMEM 1122 1143
 FT DOMAIN 1144 1170
 FT REPEAT 34 92
 FT REPEAT ? ?
 FT DOMAIN 177 367
 FT REPEAT ? ?
 FT REPEAT 423 475
 FT REPEAT 477 538
 FT REPEAT 540 599
 FT REPEAT 604 656
 FT CA_BIND 488 496
 FT CA_BIND 552 560
 FT CA_BIND 616 624
 FT SITE 472 474
 FT SITE 1146 1150
 FT DISULFID 72 81
 FT DISULFID 669 726
 FT DISULFID 778 784
 FT DISULFID 854 865
 FT DISULFID 1008 1039
 FT DISULFID 1044 1049
 FT CARBOHYD 94 94
 FT CARBOHYD 101 101
 FT CARBOHYD 332 332
 FT CARBOHYD 421 421
 FT CARBOHYD 449 449
 FT CARBOHYD 464 464
 FT CARBOHYD 588 588
 FT CARBOHYD 748 748
 FT CARBOHYD 945 945
 FT CARBOHYD 1063 1063
 FT CARBOHYD 1070 1070
 FT VARIANT 580 580
 FT VARIANT 588 588
 FT VARIANT 725 725
 SQ SEQUENCE 1170 AA; 128929 MW; ECEPFC5P2448FB1 CRC64;

Query Match 18.5%; Score 1085; DB 1; Length 1170;
 Best Local Similarity 27.6%; Pred. No. 2.1e-64;
 Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

Qy 1 FNLDTENMTQ-ENARFGGSVQL-----QSRVVVVGAPQIVAAVNRGSLYQC--DYST 54
 Db 19 YNVGLPKAKIFSGPSSEQGYAVQOFTNPKGNWLLVSPSGFPKNRGMGVKICPVDLST 78
 Qy 55 GSCBPIRLQ-----VPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGIC 107
 Db 79 TTCEKLNLTSTSMNSNTMTNMSLGLTLENVGTGGFLTQGLWAQCGSQYVTTGVC 138
 Qy 108 FLFGNLRQOKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFISTVMEQLK-- 165
 Db 139 SDVSPDF-QLRTSPAPAVQTCPSFIDVVVVCDESNSIYPWD--AVKNFLKFPVGLDIG 194
 Qy 166 KSKTLFSLMOYSEPRHFTFKFQNNPNRSLIKPITQLL-----GRTHATGIRKVVRE 221
 Db 195 PTKTQMGILQVANNPRVFNLTNFKSKD--EMIKATSQTFQYGGDLTNTFKAIQVARDT 251
 Qy 222 LFNITNGARKNAFKILILITDGEKPGDFGVYEDVPEADREGVIRYVIGV-----GDAPR 276
 Db 252 AYSTAAGCGRPGATKVMVVVTDGESH-DGSKLKAVIDCQNKDNLIFLFGIANGYLNRNLD 310
 Qy 277 SEKSRQELNTVASKPDRDHVFQINNFEALKTIONLREKIPAEIGTQSGSSSPHEMSQ 336
 Db 311 TKNLKEIKAIASIPTEHFFNVSDADLLEKAGTIGEQIFSIETGVQG-GDNFQKMSQ 369
 Qy 337 BGFSAAIT--SNGPILLSTVGSYDNAGGVPLVTSYSEKSTFINMT--RVSDMN-DAYLGYA 391
 Db 370 VGFSAEYSPQNNILMLGAVGAYDWSGTVQKTPHGLIFPSKQAPFQILQDRNHSYLGYS 429
 Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRQNTGMESNANV-----KGTQIGAFGASL 445

RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF RECEPTOR FOR LAMININS, COLLAGEN
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR A NULL MUTATION
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z29987; CAAS2877.1; -;
DR EMBL; X75427; CAAS3178.1; -;
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGI; MGI:96600; Itga2
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWFA_1
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178 INTEGRIN ALPHA-2.
FT DOMAIN 27 1129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1130 1151 POTENTIAL.
FT DOMAIN 1152 1178 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 100 FG-GAP 1.
FT REPEAT 185 375 FG-GAP 2.
FT DOMAIN 185 375 VWFA.
FT REPEAT 431 483 FG-GAP 3.
FT REPEAT 485 545 FG-GAP 4.
FT REPEAT 548 607 FG-GAP 5.
FT REPEAT 612 664 FG-GAP 6.
FT REPEAT 496 504 FG-GAP 7.
FT CA_BIND 496 504 POTENTIAL.
FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GPFK MOTIF.
FT DISULFID 80 89 BY SIMILARITY.

DB 430 VASISTGNSVHFVAGAPRANTYTCQIVLYSVN-----ENGNTVTIQSGDQIGSYFGSVL 484
QY 446 CSVDVDSNGSTDLVLICAPHYEOTR--GGQVSCVPLPGRCARWQCDVLYGEOQCPWG 503
DB 485 CAVDVNKDTITDVLVAGAPYMDLKKKEGRVYLPITKG-IILNH--QFLEQNGLENA 541
QY 504 RFGAALTVLDVNGDKLTVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KL 561
DB 542 RFGSAIAALSDINMDGDFNDVIGSPLNQNSGAVIYNHGEHM-IRLRYSKILGSDRAP 600
QY 562 SPRLQYFQSISGQDITMDGLVLTGAGQHVLLRSQPLVKVKAIMEPNPREVARNVF 621
DB 601 SSHLYQFGRSLDGGDLNGDSITDVSVGAQGVQVQLMSQSIADVSVDASTPKKI--TL 658
QY 622 ECNDQVVKGEAGEVRYCLHVQKSTRDRLEGQIQSVVYDIALD-----SGRPHSRAVEN 677
DB 659 NKNAEI-----KLKLP-----SAKRPETNNQNVAVITNIDEDQFSRVSIRGLFK 707
QY 678 ETKNSTRTQVGLTQTCR--TLKQLQPLNCIEDPSPVILRNFSI--VGTPLSAPGNL 733
DB 708 ENNERCLQKTMVISOQRCSEYIIHQEPS---DIISPLMLCMNISLENPT----- 756
QY 734 RPLVLAQRLFTALFPFKXNCGNDNICQDLSITF-----SFMSLDCLVVGGRPRENVTV 789
DB 757 NPALAEYSEIVKVFSEIPEHKDQDGGVCISDLVNLVQQLPQATQCPPIVSNQKRLTFSV 816
QY 790 TVRNGEDSTRTQVTFPPDLDSYRKVSTLQNSQORSQSWRLACESAST-EVSGALKSTS 848
DB 817 QLRNKKESAVNTEIVDFSENLF-----ASWNPVDGTEVTCQIASQKSVT 864
QY 849 CSINHPIPPENSEVFTNTPDVSQASLGNKLLKANVTSENMPNTRNKTPEQLELPVKY 908
DB 865 CNVGYPAKSKQVTFINFPDNLQ-NLQNASISPRALSESQENMADNSVNLKLSLLY 923
QY 909 AVYMTVTHGVSTKYLNFTASNTSRVMOHQYVSNLQOR-----SLPISLVPLV 958
DB 924 DAEIHIT-RSTNINFIYELSDGNVSVV-HSFE--DIGPKFIPSIKVTGSGVPVSWA--- 976
QY 959 PVRLNQTVMDRPOVTFSEN--LSSTCHTKE-----RLPSHSDFLAE- 998
DB 977 -----SVLIHPQYTKDKNPLMLVLTGHTDQAGDISCEAEINPLKIQTSSSVSFKSEN 1030
QY 999 LRKAPVWNCISAVCQRIQCDIPFGIOBEFNATLKNLSPDWYIKTSHNHLIVSTAEI- 1057
DB 1031 FRHIKELNCRATSCNIMCLWRLDQVKGEYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
QY 1058 LFNDVSFTL-----LPGQGFVRSQETKVERPE-VPNPLPLVSGSVGLLALLLITA 1110
DB 1091 TYNQIVYIENTVTIP-----LTIMKPKHEVPTGVIQSVIAGILLALLAIVA 1140
QY 1111 ALYKLGFFKQYKDM 1125
DB 1141 ILWKLGFFKRYEKM 1155

RESULT 12
ID ITA2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 43, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
GN (Collagen receptor) (VLA-2 alpha chain) (CD49b).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;

PT	DISULFID	677	734	BY SIMILARITY.	
FT	DISULFID	786	792	BY SIMILARITY.	
FT	DISULFID	862	873	BY SIMILARITY.	
FT	DISULFID	1016	1047	BY SIMILARITY.	
FT	DISULFID	1052	1057	BY SIMILARITY.	
FT	CARBOHYD	102	102	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	109	109	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	429	429	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	472	472	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1054	1054	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1078	1078	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	1178	1178	AA; 128926 MW; 1F19489C0240P465	CRC64;
Query Match					18.2%; Score 1071; DB 1; Length 1178;
Best Local Similarity					27.9%; Pred. No. 1.8e-63;
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;					
QY	1	FNLDTENMTFO-ENARGGQSVQL---	QSERVVVGAPDIEIVANORGLYQC--DYST	54	
DB	27	YVGLFGAKIFGSPSEQGYVQVQTNPQGNWLLVGS	PWSPFPPNRMGDVYKCPVDLPT	86	
QY	55	GSCPEPRLQ-----VPVEAVNMSLGLSLAATTS	PQLLACGPTVHOTCSENTVYKGLC	107	
DB	87	ATCEKLNQNSASISNVTETKNSLGLTLTENPGTGG	FLTCGLPLWAHOCGQNYATGIC	146	
QY	108	FLFGSNLRQOPQ--KFPBALRGCPQEDSDIAFLVD	SGSIIPHFRRAKPEFISTWMOQL	164	
DB	147	---SDVSPDFQFLTSFPAVQACPSL--VDVVVVCD	ESNSIYP--WEAVKFLVRFVTGL	199	
QY	165	K--KSKTLFSLMOCYSEPRHFTFKFQNNPERSLI	KITLITOLLG--RTHVATGIRKVVRE	221	
DB	200	DIGPKTKTQVALQIYANEPRIFLNFNDFTKEDVKN	QVATSETRQGGDLTNFPAIFPARDY	259	
QY	222	LFNITGARNAPKILITIDGKFDPLGDEVDVPEAD	REGVIRVIGV-----GDAPR	276	
DB	260	AYSQTSGRPGATKVMVWVVDGESH--DGSKLTVI	QCCNDDEILREGIAVLGYLRNALD	318	
QY	277	SEKSRDELNTVASKPRDRHVQINNEALQIKQNR	KEKAGTIGQIFSLGTVQG--GNFQWEMAQ	377	
DB	319	TKNLIKKEIKALASTPTERYFNFVADAEALKEA	GTIGQIFSLGTVQG--GNFQWEMAQ	377	
QY	337	EGFSA--AITSNGPLSLTVSGYDWAGVFLYTSK	ESTFINMT--RVDSDMN--DAYLGYA	391	
DB	378	VGFSAFYAPNDILMLGANGAFDWSGLTVQETSK	HPVIFPKQAFDQVLQDRNHSFLGYS	437	
QY	392	AAILNRVOSLVGAPRYQIHGLVAMFRONTGHWES	NANV-----KGTQIGAYFGASLCS	447	
DB	438	VAAISTEDGVHFFVAGAPRANTYTGQIVLYSVN	K--QGNVTVIQSHRGDQIGSYFGSVLCS	494	
QY	448	VDVDSNGSTDLVLIGAPHYEOTR--GGQVSVCP	LPGRQARQCDAVLGEGQGPWGRF	505	
DB	495	VDVDKOTITDVLVGAPTYNDLKEEGKYLTITKIL	LNQHQ---FLEGPSGTGNARF	551	
QY	506	GAALTVLGVNGDKLTVAITGARGEEDNRGAVLY	FHGTSGSGISPSHSQRIAGSKLSPR--	564	
DB	552	GSAIALSDINMDGFNDVIVGSPVENENSGAVTYN	GHQST--IRTKYSQKILSNGAFPR	610	
QY	565	-LQVFGOSLGGDLTMDGLVLTGAGQGVLLRLS	OPVLKVAIMEFNPREVARNVFE	623	
DB	611	HLQPFGRSLDGYGLNGDSITDVSIGALGVQIVL	QWSQSIADVAIEALFTP-----	660	
QY	624	NDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSW	TYDIALD----SGPSPHRAVPNET	679	
DB	661	-DKITLNLKDAKITLKLCPRAFPFAGQNNQV--	AILEFNMTLDADGHSSRVTSRGVREN	717	
QY	680	KNSTRQTVGLTQTCET--LKLQLENCEIDPSP	IVILNLSLVCTPLSAFENLRPVL	737	
DB	718	SERFLQNMVNVNQRKSEHHISIQKES---D	VNPLDLAVDISLENPGTS-----PAL	768	
QY	738	AEDAQRFLTALFPPEKKNQNDNICQDDLST	-----TFSPMSLCLVVGQPRFENVT	788	

DB	769	EAVSETVKVFSIPFYKEGSDGICISDLILDVQQL	PAIQOSP-----IVSNQNKRLTFS	823	
QY	789	VTVRENGEDSYRTQVTFPFLDLISYKVKVSTL	QNSORSWRLACESASST--EVSGALKST	847	
DB	824	VILKNRGESAYNTVLAEFSENLF-----ASF	SMYPVDGTETVEGSSQSKSV	871	
QY	848	SCSINHIFPSENSVTNITFDVDSKASLGNKLL	KANVTSENNMPTNKTE--FOLELP	905	
DB	872	TCDVGYFALKEQOVTFITPFDNLQ--NLQNA	INFQAFSESQ--ETNKADNSVSLTIP	928	
QY	906	VKYAVYVAVTSHGVSTKYLAFTASENTSVM	QHOYQVSNLQOR-----SLPISLV	955	
DB	929	LLYDAELHLT-RSTNINFIYSDDENAPSVIK--	SVEDIGPKFISLKVTAGSAPVSM	984	
QY	956	FLV-----PVRNQTVINDRPOVTF--SE	NLS	980	
DB	985	LVTHIPQYTKERKNPLLYLTGIDTQDQAGDIS	CTAIEINPLKLPHTA-----PVSFANENFR	1040	
QY	981	STCHTKERLPSSDFLAELRKAPVVGSIACV	QRCIQCDIPFGIQEHPNATLKGMLSPDW	1040	
DB	1041	---HTKE-----LDCRTTSCSNITWLKDL	HKMAEYFINVTVTRVMKRT	1080	
QY	1041	YIKTSHNHLIVSTABILFNDVSVFTLLPQCG	AFVRSQTVETKVEFEPVDPNPLPLVGVSSVG	1100	
DB	1081	FAASTFTQVQLTAAAEIDTHNPOLFVIEEN	AVTIPLMIMKTEKAEVPT--GVIIIGSIIA	1138	
QY	1101	GILLALITALAALYKLGPFKROYKDM	1125		
DB	1139	GILLALANTAGLWLGPFKROYKDM	1163		
RESULT 13					
ID	IT2A	HUMAN	STANDARD;	PRT;	1181 AA.
AC	PI7301.				
DT	01-AUG-1990	(Rel. 15, Created)			
DT	01-AUG-1990	(Rel. 15, Last sequence update)			
DE	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)				
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).				
GN	ITGA2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.			
RC	TISSUE-Endothelial cells;				
RX	MEDLINE=89308879; PubMed=2545729;				
RA	Takada Y., Hemier M.E.;				
RT	"The primary structure of the VLA-2/collagen receptor alpha 2 subunit				
RT	(platelet GP1a): homology to other integrins and the presence of a				
RT	possible collagen-binding domain."				
RL	J. Cell Biol. 109:397-407(1989).				
RN	[2]	SEQUENCE FROM N.A.			
RP	Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,				
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;				
RL	Submitted [MAY-2002] to the EMBL/GenBank/DBJ databases.				
RN	[3]	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.			
RP	MEDLINE=98019223; PubMed=9353312;				
RX	Emsley J., King S.L., Bergelson J.M., Liddington R.C.;				
RT	"Crystal structure of the I domain from integrin alpha2beta1."				
RL	J. Biol. Chem. 272:28512-28517(1997).				
RN	[4]	VARIANT HPA-5 (BR).			
RP	MEDLINE=94033762; PubMed=7901236;				
RA	Santoso S., Kaib R., Walka M., Kiefel V., Mueller-Eckhardt C.,				
RA	Newman P.J.;				
RT	"The human platelet alloantigens Br(a) and Brb are associated with a				
RT	single amino acid polymorphism on glycoprotein Ia (integrin subunit				

alpha 2).";
 J. Clin. Invest. 92:2427-2432(1993).
 [5]
 VARIANT GLU-534.
 MEDLINE=20206009; PubMed=10744142;
 Kroll H., Gardemann A., Tschert A., Haberbosch W., Santoso S.:
 "The impact of the glycoprotein Ia collagen receptor subunit A1648G
 gene polymorphism on coronary artery disease and acute myocardial
 infarction";
 Thromb. Haemost. 83:392-396(2000).
 -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT
 RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN
 COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
 EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 EXTRACELLULAR MATRIX.
 -I- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
 associates with beta-1. Interacts with HPSS.
 -I- SUBCELLULAR LOCATION: Type I membrane protein.
 -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 -I- POLYMORPHISM: Position 534 is associated with platelet-specific
 allonantigen HPA-5 (BR). HPA-5A/BR(A) has lys-534 and HPA-5B/BR(B)
 has Glu-534. HPA-5B is involved in neonatal alloimmune
 thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a
 role in coronary artery disease (CAD).
 -I- SIMILARITY: Belongs to the integrin alpha chain family.
 -I- SIMILARITY: Contains 1 VWFA domain.
 -I- SIMILARITY: Contains 7 FG-GAP repeats.
 -I- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 ENBL; X17033; CAA34894.1; --
 ENBL; AF512556; AAM34795.1; --
 PIR; A33998; A33998.
 PDB; LAOX; 25-NOV-98.
 PDB; 1DZ1; 02-AUG-01.
 Genew; HGNC:6137; ITGA2.
 MIM; 192974; --
 GO; GO:0008305; C:integrin complex; TAS.
 GO; GO:0005886; C:plasma membrane; TAS.
 GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 GO; GO:0005518; F:collagen binding; TAS.
 GO; GO:0007596; P:blood coagulation; TAS.
 GO; GO:0007160; P:cell-matrix adhesion; TAS.
 GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR002035; VWFA.
 Pfam; PF01839; FG-GAP; 3.
 Pfam; PF00357; integrin_A; 1.
 Pfam; PF00092; vwa; 1.
 SMART; SM00191; Int alpha; 5.
 SMART; SM00327; VWFA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS50234; VWFA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
 3D-structure.
 SIGNAL 1 29
 CHAIN 30 1181
 DOMAIN 30 1132
 DOMAIN 1133 1154
 TRANSMEM 1155 1181
 DOMAIN 1155 1161
 INTERACTION WITH HPSS.

FT REPEAT 45 103
 FT REPEAT ?
 FT DOMAIN 188 378
 FT VWFA.
 FT REPEAT 378 433
 FT FG-GAP 3.
 FT REPEAT 434 486
 FT FG-GAP 4.
 FT REPEAT 488 549
 FT FG-GAP 5.
 FT REPEAT 551 610
 FT FG-GAP 6.
 FT REPEAT 615 667
 FT FG-GAP 7.
 FT CA_BIND 499 507
 FT POTENTIAL.
 FT CA_BIND 563 571
 FT POTENTIAL.
 FT CA_BIND 627 635
 FT POTENTIAL.
 FT SITE 1157 1161
 FT GFPR MOTIF.
 FT BY SIMILARITY.
 FT DISULFID 83 92
 FT BY SIMILARITY.
 FT DISULFID 680 737
 FT BY SIMILARITY.
 FT DISULFID 789 795
 FT BY SIMILARITY.
 FT DISULFID 865 876
 FT BY SIMILARITY.
 FT DISULFID 1019 1050
 FT BY SIMILARITY.
 FT DISULFID 1055 1060
 FT BY SIMILARITY.
 FT CARBOHYD 105 105
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 112 112
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 343 343
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 432 432
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 475 475
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 699 699
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1057 1057
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1081 1081
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 534 534
 K -> E (IN ALLOANTIGEN HPA-5B;
 GBSNF:1801106).
 /FTIG=VAR_003977.
 FT TURN 170 171
 FT STRAND 173 180
 FT TURN 183 184
 FT HELIX 188 199
 FT TURN 201 201
 FT STRAND 204 204
 FT TURN 206 207
 FT STRAND 209 216
 FT TURN 220 224
 FT STRAND 226 228
 FT TURN 232 240
 FT HELIX 241 241
 FT TURN 242 242
 FT HELIX 262 262
 FT TURN 263 264
 FT HELIX 266 268
 FT TURN 269 269
 FT STRAND 275 282
 FT HELIX 289 291
 FT TURN 292 301
 FT STRAND 302 303
 FT TURN 304 311
 FT STRAND 313 317
 FT TURN 318 319
 FT STRAND 323 330
 FT TURN 331 332
 FT STRAND 337 340
 FT HELIX 341 344
 FT STRAND 347 353
 FT HELIX 354 362
 FT TURN 363 363
 FT TURN 363 363
 SQ SEQUENCE 1181 AA, 129295 MW, 78187BD968A94070 CRC64;
 Query Match 18.2%; Score 1068; DB 1; Length 1181;
 Best Local Similarity 26.9%; Pred No 2.9e-63;
 Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;
 QY 1 FNLDTENAMTPO-ENARGFGQSVQJ---QSGRVVVGAPQEIIVAAQROSLYOC---DYST 54
 Db 30 YNVGLPEAKIFSGPSSEQFGYAVQOQPINPKGNWLLVGSFWSGPFENRMDGVYKCPVDLST 89
 QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQOLLACGPTVHQCSENTYVKGLC 107

Db 90 ATCEKLNQTSINPVNTEKTNMGLILTRNMTGGELTGGPLWAQCCGNYTTGVC 149
 QY 108 FLFGNLRQPKQFPEALRGCEQEDSDIAFLVDGSGSIIPHDPRRAKEFISTWBEQK-- 165
 Db 150 SDISPDF-QLSASFSPATQPCFSL-IDVVVWCDESNISYFWD--AVKNFLEKFKVQGLDIG 205
 QY 166 KSKTLFSLMOYSEERFHPKFPQNNPNRSLIKETLQLLG-RHTATGIRKVVRELFN 224
 Db 206 PKTQVGLIYANNPRVFNLMYTKEMIVATSTQSYGGDLNTPGALQYARKAYS 265
 QY 225 IYNGARKNAFKILITIDCEKEGDLGVEDVITPEADREGVIRYVIGV-----GDAPRSEK 279
 Db 266 AASGGRSRTKVMVVVTDGESH-DGSMKAVIDQCNHDMILAFGLAVGLYNALDFKN 324
 QY 280 SQBELNTVASKPRPHVFNQFNFEALKTIONQIREKIPAEIGTQYGSSSPHEHNSQGF 339
 Db 325 LIKEIKATASIFTERFYFNVSDEALLEXAGTIGEQIFSIETVQG-GDNFQWMSQVGF 383
 QY 340 SAAITSNGP--LLSTVGSVDWAGGVFLYTSKEKSTFINNT--RVDSDMN-DAYLGYAAAI 394
 Db 384 SADYSSQNDILMLGAVGAGWSGTIVKTSHGHLFPKQAFDQILQDRNHSYLGYSVAA 443
 QY 395 ILNRVQSVLVGAPRYQHOHTGLVAMFRONTGMESNANV-----KGTQIGAVFGSLCSV 448
 Db 444 ISTGSTHFAVAGPRANYTGQIVLYSVN-----ENGNITVIOAHRGDOQIGSYFGSVLCV 498
 QY 449 DVDSNGSTDLVLIGAPHYEQTR--GGQVSVCPFLPRGQARWQCDVAVLGEQ-QPWGRFG 506
 Db 499 DVDKTITDVLVGAAPMYMSDLKEGRVYLTIKKILGQHQ-----FLEGPEGIENTRFG 555
 QY 507 AALTVLGVNKGKLTDAVIGAFCEEDNRGAVLPHGTSGSGISPSHSQRIAGS--KLSPR 564
 Db 556 SAIAALSDINDGCFNDVIVGSPLENQNGAVYINGHQT-IRTKYSQKILGSDGAFRSH 614
 QY 565 LOYFGSLGGQDLTHDGLVDITVQNGCHVILLRSQPLVRVKAIMEFREVARNVFCN 624
 Db 615 LOYFGSLDGYDGLNGDSITDSIGAFQVQVQLWSQSIADVAIBASFTPEKI--TLVNKN 672
 QY 625 DQVWKGKZGAEVRVCLHVQKSTRDRRECGIQSVVTVYDLALD-----SGRPHSRVAFNETK 680
 Db 673 AQII-----LKLCE-----SAKFRTKQNNQVAIVYNTILDADGSSRVTSRGLFKENN 721
 QY 681 NSTRTQVGLTQTC--BTLKQLPNCIEDPVSPVILNLSVLTGTPLSAGNLRPVLA 738
 Db 722 ERCLQNMVYNAQSCPEHIIYIPEPS--DVVNSLDLVDLSLENPGTS-----PALE 772
 QY 739 EDAQRLFTALFPFKKNGDNICODDLSTF-----SEMSLDCLVWGGPREFVNTVVRND 794
 Db 773 AYSETAKVSIIPHKDCGEGDGLCISDLVLRQIPAAQOPFIVSNQNKRLFTSVTLKNK 832
 QY 795 GEDSVYRTQVTFPPDLVSRKXSTLQNSQSRWRLACESAST-EVSGALKSTSCSINH 853
 Db 833 RESAVNTGIWDPSENLF-----ASFSLPVDGTETVTCQVAASQKSVACDVGY 880
 QY 854 PIPPENSEVTNITDQVDSKASIGNKLLKANVTSENMMRTKTEFQELPVPKAVVAV 913
 Db 881 PALKRQQTFTTINFDNFIQ-NIQNSQSLSPQALSESQENKADNLVNLKIPLLVDAEI- 938
 QY 914 VTSHGVSRTKYLNPASENTSRYNQHOVQVSNLQGR-----SLPISLVPLV----- 958
 Db 939 ---HLRSTNINFPYELSSDGNVPSIVHSPEDVGKPIFSLKVTGSGVPVSMATVHIHPQ 995
 QY 959 -----PVLNQTIVNDRQVTF-SENLSSTCHTKER 988
 Db 996 YTKRKNPLMYLTVGVQTDKAGDISCNADINPLTKGTQ-----SSSVSPKSNFR---HTKE- 1047
 QY 989 LPSSHDFLAELRKAPVYVNSIAVCQRIQCDIPFGIQEENFATLKNLSFDNVIKTSNKH 1048
 Db 1048 -----LNCRATSCSNVTCWLKDVHMKGEYFVNVVTTIRIWNNGTFASSTFQT 1091
 QY 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRQCTETKVEPPEVNP-----LP--LIVGSSVG 1100
 Db 1092 VQLTAAARINTYNPFIYVI-----EDNTVTIPLIMKPDKAEVPTGTGIIGSTIA 1141

QY 1101 GLLLLALITAAALYKLGFFKQYKDM 1125
 Db 1142 GILLALVALKLGFFKQYKDM 1166

RESULT 14

ITAG_HUMAN STANDARD; PRT; 1167 AA.
 AC 075578; Q9UHZ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-10 precursor.
 GN ITGA10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Articular chondrocytes;
 RX MEDLINE=98352078; PubMed=9685391;
 RA Camper L., Hellman U., Lundgren-Akerlund B.;
 RA "Isolation, cloning, and sequence analysis of the integrin subunit
 alpha10, a beta1-associated collagen binding integrin expressed on
 chondrocytes.";
 RL J. Biol. Chem. 273:20383-20389(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endothelial cells, and Heart;
 RX MEDLINE=20169197; PubMed=10702680;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
 RA Wang S.-X., Langley R., Krissansen G.W.;
 RT "The integrin alpha10 subunit: expression pattern, partial gene
 structure, and chromosomal localization.";
 RL Cytogenet. Cell Genet. 87:238-244(1999).
 CC -1- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
 CC ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest expression in
 CC muscle and heart. Found in articular cartilage.
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF074015; AAC31952.1; --
 CC EMBL; AF112345; AAF21944.1; --
 CC EMBL; AF172723; AAF61638.1; --
 CC HSSP; P17301; 1A0X.
 CC Genew; HGNC:6135; ITGA10.
 CC MIM; 604042; --
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.

DR SMART; SMO0191; Int_alpha: 4.
DR SMART; SMO0327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
DR PROSITE; PS00234; VWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1167 INTEGRIN ALPHA-10.
FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1123 1145 POTENTIAL.
FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT 38 97 FG-GAP 1.
FT REPEAT 39 98 FG-GAP 2.
FT REPEAT 167 350 VWA.
FT REPEAT 365 427 FG-GAP 3.
FT REPEAT 428 482 FG-GAP 4.
FT REPEAT 483 545 FG-GAP 5.
FT REPEAT 546 605 FG-GAP 6.
FT REPEAT 608 660 FG-GAP 7.
FT DOMAIN 1134 1140 POLY-LEU.
FT CA_BIND 494 502 POTENTIAL.
FT CA_BIND 558 566 POTENTIAL.
FT DISULFID 76 86 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 681 736 BY SIMILARITY.
FT DISULFID 789 795 BY SIMILARITY.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 844 844 I -> L (IN REF. 2).
FT CONFLICT 909 909 G -> V (IN REF. 2).
FT CONFLICT 926 926 E -> D (IN REF. 2).
SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AE80 CRC64;

Query Match 18.1%; Score 1064.5; DB 1; Length 1167;
Best local similarity 29.0%; Pred. No. 4.9e-63;
Matches 353; Conservative 201; Mismatches 493; Indels 171; Gaps 44;

QY 1 ENLDTENAMTFQENARG-FQGSVVOLOGSR---VVVGAPOEIVAAHORGSLYQC----- 50
DB 23 FALDEHPLPFGPPPEAFEGYVLOHVGGRWMLVGAFWDPGSDRRGDVTRCPVGGAH 82
QY 51 -----DYSTG-SCBPRLQVPEAVNMSLGLSLAATSPPOLLAGCGTTHVQTCEE 99
DB 83 NAPCAKGLHGLDYOLGNSSHP-----AVNMHLGMSHLETDGCGFMACAPLWSRACGS 134
QY 100 NTYVKGCLPLFGSNLFCQPKPEALRGCPQSDSDIAFLVDGSGSII PHDFRAKEFLST 159
DB 135 SVFSSGICARVDASFOGSLAPTAOR-CPTY-MDVIVLDGNSIYP--WSEVOTFLR 190
QY 160 VMEQL--KSKSTLFLSMQYSSEPHIHTFKKEFQNNPNRSLIKPITOLLGR-THTATGIR 216
DB 191 LVKFLIDPEQIOGLVQYGESPVHWSLGDFTKEEVVRAAKNLSRREGRETKTAQAIM 250
QY 217 KVVRELFNITNGARKAFILILITGEKEF-GDPLGYEDVIEADREGVIRVIGV-GDA 274
DB 251 VACTGFSOSHGRPEARLLVVVVDGESHDEELPAALKACEAGR--VTRYGIAVLGHY 308
QY 275 FRSEKS----ROBLNIVASKPPRDHVFQINNFEALKTIONQLREKIFAIEGTQTGSSSSF 330
DB 309 LRRQDPSSFLREIRTIASDPDRFFNFVNTDEAALTDIVDALGDRIFLEGSHAENESSF 368
QY 331 EHEMSQEGSAITSGNPLLSVTGSDWAGGVFLYTSKEKSFINKTRVDS-----DMND 385
DB 385

369 GLEMSQIGFSTHRLKDGILFGWGAIDWGGSVLMLEGHRLPPRMALEDEFPFPAQONHA 428
386 AYLGYA-AAIILNRNVQSLVILGAPRYQHIGLVAMFR-ONTQGWESNANVKGTQIGAYFGA 443
429 AYLGYSVSSMLRGERRLLFGAPFRERGVKVIAPOLKDKGAVRVAQSLQEQIGSYFGS 488
444 SECSVDVDSNGSTDLVILGAPHY--EQTREGQYVSVCLPREGQARQWQCDVLYGEQGP 501
489 ELCPUDTDRDGTUULLVAAPFWLGPQNKETGRVTVILV--GQSSLLTLOGTLQEPFPQD 546
502 WGRFGAALTVLGDVNGDKLTDVAICGPEEDNRGAVILFHGTSGSGISPSHSORIASKL 561
547 -ARTGFAMGALPDLMQDGFADVAVGAPLEDHGQALYLHGTQ--SGVRPHPAQRIAAASM 604
562 SPRLQYFQGSLSGGQDLTNDGLVBLTVCAQGHVLLRSQVLRVKAIMEFNRERVARNVF 621
605 PHALSYFGRSYDGRILDGDDLDVAVAGAAILLSSRPVHLTPSLVETFEQALISVVQR 664
622 ECNDQVVKGEAG--EVRVCLHVKSTRDRREGQISVVVYDLDLSDGRSHSRAVNET 679
665 DCR---RGQEA VGLTAALCFQVTSRTPTGRWDH---QPYMEFTASLDEWTAAGAAAFDGS 718
680 --KNSTRQTOVLGUTQTCETLKLQPNCEIDPVSPIVLRNLSVLTGTPSLAFGNLRVIL 737
719 CQRLSPRLRLSVG-NVTCEQLHFHVLQ-TSDYLRLPVALTVTFALDNTTKFG-----PVL 771
738 AEDAQRLPTALPPKKNCGNDNICDDLSITFSFMSLDC-----LVVGGPREFNAVTV 789
772 NEGSPTSLOKLVPFSKDCGPDNECVTDLVQ---VWMDIRGRKAPFVVRGRRKVLVST 828
790 TVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRSLACLASASSTEVSGALKSTSC 849
829 TLENKENAVNTSLIIFSRNL---HLASLTPQR-ESPIKVECAAPSA-----HARLC 877
850 SINHPPIPEENSEVTNITFDVDSKASLG--NKL---LKANVTSENNMPRTNKTEFQ 901
878 SVGHPVGTGAKVTLLEFEFSCSLLSQVFGKLTASSDSLRNGTLQENTAQT----- 931
902 LELPVKAVVMVTSHGVSSTKYLNFMTASENTSRVQHGY-----QVSNLG 946
932 -----SAVQYEPH-----LLFSSSEITLHRYEVHPYGLPVGPGPEPKTLRVQNLG 978
947 ---QESLPLSLVFLVP-----VRLNQTVIWDPRQVTFSENLSSCTCHKERLPUSD 994
979 CVYVSGLLISA--LLPAVAHGNFTLSLSQVI-----TNASCIVQMLTEPFGPPV 1027
995 FLAHLKAPVYVNCIAVCQRIQCDIPFGIOEENFATLKGNSLFDWYIKTSHNHLIYST 1054
1028 HPEELQHTNRLNGSNTQCVVRCHLGQAKCTEVSGLRLVHNEFPFRAXFKSLTVVST 1087
1055 AEILLNDSVFTLLPOGAFVPSQTEKVEPEVPNPPLIVGSSVGGLLILALITAAVYK 1114
1088 FELGTEGSLVQLTEASRWSSESLLEV-VQTPILLISLWILIGSVLGGLLALLVFCLEWK 1146
1115 LGFF-----KROYK 1123
1147 LGFFAHKKIPEEKREEK 1164

RESULT 15
ITAL RAT
ID ITAL RAT STANDARD; PRT; 1180 AA.
AC P19614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

RN SEQUENCE FROM N.A.
 RP MEDLINE=90338125; PubMed=2380249;
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
 RA Esch P., Carbonetto S., Reichardt L.F.,
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
 RT for laminin and collagen.";
 RL J. Cell Biol. 111:709-720(1990).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
 RA MEDLINE=99313197; PubMed=10386626;
 RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,
 RA Gortals P.J., Karpusas M.,
 RT "Crystal structure of the alpha1beta1 integrin I-domain: insights into
 RT integrin I-domain function.";
 RL FEBS Lett. 452:379-385(1999).
 CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
 CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
 CC E-R IN COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X52140; CAA36384.1; -;
 DR PIR; A35854; A35854.
 DR PDB; 1CK4; 03-MAY-00.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP_3;
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWFA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
 FT SIGNAL 1 28
 FT CHAIN 29 1180 INTEGRIN ALPHA-1.
 FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1143 1165 POTENTIAL.
 FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 44 103 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT REPEAT 175 388 VWFA.
 FT REPEAT 377 432 FG-GAP 3.
 FT REPEAT 433 484 FG-GAP 4.
 FT REPEAT 485 565 FG-GAP 5.
 FT REPEAT 567 626 FG-GAP 6.
 FT REPEAT 629 681 FG-GAP 7.
 FT CA_BIND 497 505 POTENTIAL.
 FT CA_BIND 579 587 POTENTIAL.
 FT CA_BIND 641 649 POTENTIAL.
 FT SITE 1168 1172 GPFKR MOTIF.
 FT DISULFID 82 92 BY SIMILARITY.
 FT DISULFID 687 696 BY SIMILARITY.
 FT DISULFID 702 755 BY SIMILARITY.
 FT DISULFID 807 813 BY SIMILARITY.
 FT DISULFID 877 885 BY SIMILARITY.
 FT DISULFID 1029 1062 BY SIMILARITY.

FT	DISULFID	1066	1073	BY SIMILARITY.
FT	CARBOHYD	100	100	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	105	105	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	217	217	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	317	317	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	341	341	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	402	402	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	418	418	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	459	459	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	531	531	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	698	698	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	747	747	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	779	779	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	820	820	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	839	839	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	882	882	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	907	907	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	938	938	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	965	965	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	973	973	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1007	1007	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1084	1084	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1103	1103	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1114	1114	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1180 AA;	130808 MW;	8E5DA23E02362EB4 CRC64;

Query Match 18.1%; Score 1062; DB 1; Length 1180;
 Best Local Similarity 27.4%; Pred. No. 7.4e-63;
 Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;
 QY 1 FNLPTENAMTFOENARG-FGQSVVQL---QGSRRVVGAGQEIIVAAVNGRSLVQCDYSTGS 56
 DB 29 FNVDKNSMSPGSDVEDMFVTVQVYENBEGKWLIGSPVGVQPKARTGDVYKCPVGRER 88
 QY 57 CEP-IRLQVPEA-----VNMSLGLSLAATSPQLLACGTPVHQTCSNTYVYKGL 106
 DB 89 AMPCCVCLDLPVNTSPINVTKEIKENMTFGSTL-VTNPMGGFLACGLFYARCGHLHYTGI 147
 QY 107 CFLFGSNLRQOPKFPFALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPFTVMBQLK- 165
 DB 148 CSDVSPTFQVNSFAP--VQECSTQ-LDIIVLDGNSIYP--WESVIAFLNDLLKRMDI 202
 QY 166 -KSKTLFSLMOYSEEPRIHFTFKFQNNPNSLIKPTQLLG-RTHATGIRKVVREL 223
 DB 203 GPKQTQVGVQYGENVTHEFNLNKYSTEBVLVAANKIGQGGQCTWTALGIDTARKEAF 262
 QY 224 NITNGARKNAFKILILITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPR-----SE 278
 DB 263 TEARGARGVKVWVIVTDGESH-DNYELKQVQDCEDENIQRFSAIILGHYNRGNLSTE 321
 QY 279 KSRQELNVTASKPRDRHVQINNFEALKTQNLQREKIPAIETGTCGSSSFEHMSQEG 338
 DB 322 KFVEEIKSIASBPTERHFVNSDELALVTIVKALGERIFALEATADQSAASFEMMSQTG 381
 QY 339 FSAAITSNGLPLSTVGSVDWAGGVFLVTSKEKSTFINNT--RVDSMDND---AYLGYAAA 393
 DB 382 FSAHYSQDWMLGAYDNGTVMQKQWVPHNTTQTETPAKWEPLASVIGITVN 441
 QY 394 IILERNVQSLVLAGPRYOHIGLVAMFRONTQWESNANVKQTQIGAYFGALCSVDVDSN 453
 DB 442 SATIPGDVLYIAGQPRYNHTQVVIYKMGEDNINILQTLGGEQIGSYFGSLVTTIDDKD 501
 QY 454 GSTDLVLIGAPHY-----YEQTR--GGQVSVCLPRGQARWQCDAVLGE 497
 DB 502 SYTDLLLVGAPMYNGTEKEEGKYVTVAVNQTRFYQMSLEPIRQTCCSLKXDNCSCTKEN 561
 QY 498 QGQPMWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLPFGHTSGSGISPHSQRI 556
 DB 562 KNEFGARFGTAIAAVKDLNVGDFNDVVIGAPLDDHAGAVIYHG-SGKTIREAYAQR 620
 QY 557 AGSKLSPLQVFGQSLSGGQDLTMDGLVDLTVGAGQHVILLRSQPLVLEVKALMEPREV 616

Db 621 PSGDGKTLKFFGQSIHGMDLNGDGLTDVTIGLGAALFWARDVAVVVKVTMPEPNKV 680
QY 617 ARNVFECNDQVVKREAG--EVRVCLHVQ-KSTRDLREGQIQSVVYDLDLDSGRPHSR 673
Db 681 N:QKKNCR---VEGKETVCINATMCFHVKLSKEDSIYEADLQ-----YRVTLDSLRQISR 733
QY 674 AVFNET-----KNSTRQTVGLTQTCTETKLQLPNCI-----EDFVSPVILRL 718
Db 734 SFSGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVRVTL 776
QY 719 NFSLVGTPLSAGNLRPVLAEDAQRDLFTALFFPKXCGNDNICQDDLSITTFKSLDCLV 778
Db 777 DFNLT-DPENG-----PVLDDALPNSVHEHIFPAKDCGNKERCISDLTLNVSTTEKSLLI 830
QY 779 VGGPRE-FNVVTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNSORSQSWRLACESASS 837
Db 831 VFSQHDKNVSLTVKNGDSANTRTVVQHSNLI FSGIEEIQD-----SCESN-- 880
QY 838 TEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLL-LKANVTSENMPRTN 896
Db 881 -----QNTICRVGYPLRAGETVTFKLIQFNTSHLSENAIHL SATSDSEEPLESIN 933
QY 897 KTEFOLELPKYAV---YMWVTSHGVST-----KYLNTASENTSRVMOHQYQVSNL 945
Db 934 DNEVNISIFVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFTIRKR 991
QY 946 GORSLP---ISLVF-----LVPVRLNOTVIND-----RP-----Q 972
Db 992 GHFPMPELQLSISFPNLTADGYVLYPIG-----WSSSDNVNCRPSRLEDPFGINSQKK 1045
QY 973 VIFS-----ENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCRIQCDI--- 1019
Db 1046 MTISKSEVLKRTIQDCSSTC-----GVATITCSLPSDLSQVNVSL 1088
QY 1020 ---PFFGIQEEP---NATLKGNLSFDWYIKTSHNHLIVSTABILLFNDVFTLLPQQGAF 1073
Db 1089 LMKPTF-IRAHFSSNLTLRGELK-----SENSSLTLSSN----- 1123
QY 1074 VRSQETKVEPEVPNPPLP--IVGSSVGGLILALITAAKYKLGEPKQYKMMSE 1128
Db 1124 RKZELAIQISKDGLPGRVPLWILLSAPAGLLLELLMLLILALMKIGFPKPKKKKKMEK 1180

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds
(without alignments)
9084.693 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLTENANTQENARGFGQ.....FKRQVKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4381	74.6	1151	11 Q9J130	Q9J130 rattus norv
2	3901.5	66.4	1036	11 Q8CA73	Q8CA73 mus musculu
3	3809.5	64.8	920	6 Q28984	Q28984 sus scrofa
4	3460	59.2	1169	4 Q8IVA6	Q8IVA6 homo sapien
5	3320.5	56.5	1169	11 Q9QXH4	Q9QXH4 mus musculu
6	3240.5	55.2	1161	11 Q9QYB7	Q9QYB7 rattus norv
7	1534.5	26.1	1161	11 Q9RTV4	Q9RTV4 mus musculu
8	1524	25.9	1160	11 Q9R200	Q9R200 mus musculu
9	1409	24.0	1196	13 Q98TF1	Q98TF1 cyprinus ca
10	1359.5	23.1	1086	4 Q96H31	Q96H31 homo sapien
11	1358.5	23.1	1187	13 Q98TF0	Q98TF0 cyprinus ca
12	1278	21.7	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1167.5	19.9	1167	11 Q88340	Q88340 rattus norv
14	1123	19.1	1167	11 Q88341	Q88341 rattus norv
15	1053	17.9	1171	13 Q42094	Q42094 gallus gall
16	1052.5	17.9	1188	11 Q7TQC3	Q7TQC3 mus musculu

17	1045	17.8	1038	11 Q8BS01	Q8BS01 mus musculu
18	1018.5	17.3	1160	6 Q8MKF4	Q8MKF4 felis silve
19	1018	17.3	895	11 Q9WUF8	Q9WUF8 mus sp. itg
20	865	14.7	348	4 Q8TES5	Q8TES5 homo sapien
21	850	14.5	1332	5 Q9BPQ8	Q9BPQ8 halocynthia
22	808	13.8	205	11 Q63001	Q63001 rattus norv
23	753.5	12.8	780	13 Q06271	Q06271 xenopus lae
24	738	12.6	823	4 Q8WYI8	Q8WYI8 homo sapien
25	686.5	11.7	823	11 Q8CB84	Q8CB84 mus musculu
26	669	11.4	1032	11 Q61989	Q61989 mus musculu
27	643	10.9	1036	11 Q91YD5	Q91YD5 mus musculu
28	642.5	10.9	1033	6 Q9BGU3	Q9BGU3 bos taurus
29	627.5	10.7	1474	5 Q86GB7	Q86GB7 pseudoplusi
30	623.5	10.6	257	11 Q8C270	Q8C270 mus musculu
31	619.5	10.5	1041	5 Q9UB90	Q9UB90 lytechinus
32	614.5	10.5	1041	5 Q76378	Q76378 lytechinus
33	579.5	9.9	1054	5 Q9U6S1	Q9U6S1 strongyloce
34	578.5	9.8	1034	13 Q98TF7	Q98TF7 gallus gall
35	555.5	9.5	1053	11 Q80TP5	Q80TP5 mus musculu
36	550	9.4	1033	13 Q42598	Q42598 xenopus lae
37	546	9.3	1036	6 Q7YRP8	Q7YRP8 equus cabal
38	534	9.1	1016	13 Q91779	Q91779 xenopus lae
39	530	9.0	974	11 Q924W2	Q924W2 rattus norv
40	529	9.0	1073	11 Q8CC06	Q8CC06 mus musculu
41	528	9.0	1119	5 Q86GS8	Q86GS8 pseudoplusi
42	526	9.0	1047	6 Q9MZB6	Q9MZB6 bos taurus
43	525.5	8.9	1007	6 Q9GK48	Q9GK48 bos taurus
44	522.5	8.9	1132	11 Q80Z18	Q80Z18 mus musculu
45	512.5	8.7	1034	6 Q9TUN4	Q9TUN4 oryctolagus

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrila K. Jr.;
RT Cloning of the rat CD11b cDNA sequence.
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA_1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 11; Length 1151;
Best Local Similarity 72.7%; Pred. No. 1.3e-315;
Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;

361 GVFLYTSKESKSTFINMTRVSDMNDAYLGYAAALILRNVRQSLVGLGAPRYQHIGLVAMER 420
377 GAFYTSKDKVTFINTFVSDMNDAYLGYASAVILRNVRQSLVGLGAPRYQHIGLVAMER 436
421 QNTGWESNANVKTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHTEQTRGGQSVCP 480
437 ENFGTWEPHTSING----- 450
481 PRGQARQCDNAVLYGEGQPGWRFGAALTGLVDVNGDKLTDVAIGAPGBEDNRGAVYLF 540
451 ----- 450
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVILLRSQ 600
451 -----SQRIIGAHFSPGLQYFGQSLSGQDLTMDGLMDLAVGAQGHVILLRAQ 498
601 PVLRVKATMEFNPREVARNVECDNVVKGKAGSVRVLHVOKSTRDLREGOIQSVWT 660
499 PVLRELEATWESPKKVAASVACQOVUKNDKAGEVRVCLVRVNTKDRLEGDIQSTVT 558
661 YDLALDSGRPSRAVFNETKSTRTOVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
559 YDLALDPGSRIRAFDETKNTRERTQVFLGMQKCELTLLPDCVDDSVSPILRLNY 618
721 SLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSKMSLDCLVVG 780
619 TLVGEPLSFGNLRPVLAEADAQRFFTAMPPEKNCNDNI CODDLSITMSANGDITLVVG 678
781 GPREENVTVVRNDEGDSYRTQVTFEFPFLDLISYKRVSTLQORSQSWRL-ACESASSTE 839
679 GPQDNFMSVTLNDEGDSYGTQVTVYPSGLSYKDSASQPLTKKPFVVKPAESSSSSE 738
840 VSGALKSTSCSINHPIFENSVTNITFDVDSKASLGNKLLKANKANTSENNMERTNTE 899
739 GHGALKSTWNIHPIFANSEVTNITFDVDSHASFGNKLKLLKAIIVASENNMERTNTE 798
900 FOLEPVPKAVVMTSHGVSTKYLNFETASENTSRVMOHOYQVNSLQORSIPISLVFLVP 959
799 FOLEPVPKATVMTSDSSRYLNFETASENTSVIHOYQVNSLQORSIPISLVFLVP 858
960 VRLNQTIVWRDQVTFSENLSCTHCKRLPGHSDFLAELRKAQVNVNCSIAVCQRIQDI 1019
859 VQINNVTVMDHQPVPSQNLSSACHTEQSPHNSFRDQLERTPVLNCSVAVCKRIQCDL 918
1020 PFGQOEBENATLKNLSFDWIKTSHNLLIVSTAEIILFNDSVFTLLPGQCAFVRSOTE 1079
919 PSFNTQOEIFNVTLKNLSFDWIKTSHGELLVSSTEILFNDSAFALLPQGESYVRSKTE 978
1080 TKVEPEVENPPLIVGSSVGGILLIALITALYKLGFFKQYKDMMSGCGPPGAPQ 1137
979 TKVEPEVENPPLIVGSSIGGIVLIALITALYKLGFFKQYKDMMSGCGPPGAPQ 1036

RESULT 3
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DE 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN CD11B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCES FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U40072; AAB16869.1; --
DR HSP; P11215; 1BHQ.

GO: GO:0008305; C:integrin complex; IEA.
GO: GO:0004895; P:cell adhesion receptor activity; IEA.
GO: GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF A.
DR Pfam: PF01839; EG-GAP; 3.
DR Pfam: PF00092; Vwf; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS0234; VWF A; 1.
FT NON_TER 1 1
FT NON_TER 920 920
SQ SEQUENCE 920 AA; 102440 MW; E96CCS1E350DD5AC CRC64;

Query Match 64.8%; Score 3809.5; DB 6; Length 920;
Best Local Similarity 78.9%; Pred. No. 2.3e-273;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;

QY 118 PQFPEALRCPCQEDSDIAFLVDGSGSIIPHDERRAKEFTISTVMEQLKKSKTLFSLAQYS 177
DB 1 PQFPEALRCPCQEDSDIAFLVDGSGSIINFLDPRKKEFTVYMGQFQKSKTLFALQYS 60
QY 178 EEFRIHFTFXEQNNPNRSLIKPITOLLGRTHATGIRKVVRELFWITNGARKNAKFL 237
DB 61 EDFYTHFTFNDKRNPSKLLVRPIROLLGRTHATGIRKVVRELPHSKSGARENAKFL 120
QY 238 ILITDCKFGDPLYEDVIEADREGVIRYVIGVDAFRSEKSRQELNIVASKEPRDHVF 297
DB 121 WITDSEKFGDPLYEDVIEADREGVIRYVIGVDAFRSEKSRQELNIVASKEPRDHVF 180
QY 298 QINNFEALKTIQNLREKIFAIECTQTGSSSSSPHEMSQEGFSAITNSGPLLSTVGSYD 357
DB 181 QVNNFEAVKTIQNLQEKTAIECTQTGSSSPHEMSQEGFSAITNSGPLLSTVGSYD 240
QY 358 WAGGVFLYTSKESKSTFINMTRVSDMNDAYLGYAAALILRNVRQSLVGLGAPRYQHIGLVA 417
DB 241 WAGGAFLHMPKORVIFINTFVSDMNDAYLGYAVEVILRNQASLVGLGAPRYQHTGLVV 300
QY 418 MFRQNTGWESNANVKTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYHTEQTRGGQSV 477
DB 301 MFKNSGANWEKADIKSOIGSYFGSLSVDVNRDGSDDLVLIGAPHYHTEQTRGGQSV 360
QY 478 CPLPRGQARQCDNAVLYGEGQPGWRFGAALTGLVDVNGDKLTDVAIGAPGBEDNRGAV 537
DB 361 CPLPQGR-RAKQCRVILCGEQGPWSRFGAALTGLVDVNGDKLTDVAIGAPGBEDNRGAV 419
QY 538 YLFHCTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVILL 597
DB 420 YLFHCTSELGSPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLMDLAVGAQGHVILL 479
QY 598 RSQPVLRVYKATMEFNPREVARNVECDNVVKGKAGSVRVLHVOKSTRDLREGOIQS 657
DB 480 RSQPVLRVYKATMEFNPREVARNVECDNVVKGKAGSVRVLHVOKSTRDLREGOIQS 539
QY 658 VVTVDLALDSGRPSRAVFNETKSTRTOVLGTQTCETLKLQLPNCIEDPVPVILRL 717
DB 540 IITVDLALDPGRPHPRVAFBETKNTNRTQTVLGSRAKCEHALMLPDCVEDSVTPVILR 599
QY 718 LNFSLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSKMSLDCL 777
DB 600 LNFSLVGTPLSAPGNLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSKMSLDCL 659
QY 778 VVGGPREFNVTVVRNDEGDSYRTQVTFEFPFLDLISYKRVSTLQORSQSWRLACESAS 837
DB 660 VVGGPREFNVTVVRNDEGDSYRTQVTFEFPFLDLISYKRVSTLQORSQSWRLACESAS 719
QY 838 TEVSGALKSTSCSINHPIFENSSEVTNITFDVDSKASLGNKLLKANKANTSENNMERTN 897
DB 720 TEESTALKSTSCSINHPIFENSSEVTNITFDVDSKASLGNKLLKANKANTSENNMERTN 779
QY 898 TEFQLELPVKAVVMTSHGVSTKYLNFETASENTSRVMOHOYQVNSLQORSIPISLVFL 957

```

Db 780 TEFQLELPVYAVVVVVTSLVSTKYFNFTASEKTRHVIHQYQFNNGQRLKLPISVVWF 839
Qy 958 VVRLNCTVWDPRPOVTFPSNLSSTCHTKERLPSDSDFLAELRKAPVNVNCSIAVCORIQ 1017
Db 840 VVRLNCTVWDPRPOVTFPSNLSSTCHTKERLPSDSDFLAELRKAPVNVNCSIAVCORIQ 899
Qy 1018 DIPFGIOEENATLKNLSF 1038
Db 900 DIPFGIOEENATLKNLSF 920

RESULT 4
Q8IVAG PRELIMINARY; PRT; 1169 AA.
AC Q8IVAG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Scrausberg K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; 17B484FEFC79EB6 CRC64;

Query Match 59.2%; Score 3480; DB 4; Length 1169;
Best Local Similarity 61.2%; Pident. No. 1e-248;
Matches 691; Conservative 141; Mismatches 291; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFQENAFGQSVVQVQSGSVVVGAPQEIIVANQKSLYQDYSTGSCBPI 60
Db 20 FNLDTTELTAFRVDSAGFGDSVVQVANSVVVVGAPQKI TAANQGLYQCGYSTGACBPI 79

Qy 61 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHOTCSNTVYVKGICFLFGSNLRQOPK 120
Db 80 GLQVPPEAVNMSLGLSLAATTSPQALLACGPTVHOTCSNTVYVKGICFLFGSNLRQOPK 137

Qy 121 FPEALGCPQEDSDIAFLVDGSGSIIPDFRAKEPFI STWMBQLKSKTFLFSLMQYSEEP 180
Db 138 LPVSRQECPRQEDIVFLDGGSGSISSRNFAFMFRAVISOQRPSPQPSLMQFSNKP 197

Qy 181 RHPTFKQNNPNSRLKPI TOLLGRHTATGIRKVVRELPNTNGARKNAFKILILI 240
Db 198 QHPTFKQNNPNSRLKPI TOLLGRHTATGIRKVVRELPNTNGARKNAFKILILI 257

Qy 241 TDGEKFGDPLGYEDVPEADRGVIRYVIGVGDAFRSEKSRQELNVAASKPPDRHVQIN 300
Db 258 TDGKKGDSLDYKDVIPWADAAGIIRYAIGVGLAFQNRNSWKELDIAKPSQEHIFKVE 317

Qy 301 NFEALKTIONLREKIPATECTQTSSSPEHEMSQEGSAITNSGILLSTVGSYDNAG 360
Db 318 DFDALQIQONLQKIKFATEGTETTTSSSFELEMAQEGFSAVFTPDGVPVGVGSGFTWSG 377
```

```

Qy 361 GVFLYTSKEKSTFTNMTTRVDSMDMDAYLGAAAIILNRNVOSLVLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPNMPTFTNMQSDVMDSDYLGSTELALAKGVQSLVGLGAPRYOHTKAVIF 437
Qy 421 QNTGWESNANVKCTOIGAYFGASLCSVDVDSNGSTDLVILIGAPHYEQTRGGQSVQCP 480
Db 438 QVSWRKAEVCTOIGAYFGASLCSVDVDSNGSTDLVILIGAPHYEQTRGGQSVQCP 497
Qy 481 PRQARAWQCDVILYGGQGPWGFAGALTVLGVNGBDKLTDVAIGAPBEDNEGAYVLP 540
Db 498 PRGWR-RWNCDAVLYGQGHGPGWGFAGALTVLGVNGBDKLTDVVGAPGKENGAYVLP 556
Qy 541 HGTSGSISPSHSORIASGLPELOYPGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPISPSHSORIASGLPELOYPGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRYKAIMENPREVARNYFECNDQVVKGEAGEVRVCLHVQKSTRDLREGEQIQSVT 660
Db 617 PVLWVGVMQPIPAEIPSAFECEQVQVSEGTIVQSNICLYIDIKRKNLLGSRDLQSSVT 676
Qy 661 YDLALDGRPHSRAVNETKNSFRQVGLTCTCETLKLQPNCTEDPVSPTVLRNF 720
Db 677 LDALDPGLRSPRATFOETKNSRVRVILGLKACENFNLLPSCVEDSVTPITLRNF 736
Qy 721 SLVGTPLSAFNGNLPVLAEDAQRLETFALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 TLVCKPLAFENLRPMLAADAQRYFTASLPPEKNCAGADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPRENTVTVVNDGEDSYRQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 SNLEINAEVVMVNDGEDSYGTTVTFSPAGLSYVVAEGQKQGLRSLHLTCDSPAVG-- 854
Qy 841 SGALKSTSCSINHPIFPENSEVETNITPFDVDSKASLGNKLLKANVTSENNMPRTNTEP 900
Db 855 SQGTWSTSCRINHPIFRGGAQITPLATFDVSPKAVLGRLLLTANVSNNTPRTSKTF 914
Qy 901 QLELPVYAVVTVSHGVSTKYINFTAS-ENTSRVMQHQVQVSNLQORSPLSLVLP 959
Db 915 QLELPVYAVVTVSHGVSTKYINFTAS-ENTSRVMQHQVQVSNLQORSPLSLVLP 974
Qy 960 VRLNCTVWDPRPOVTFPSNLSSTCHTKERLPSDSDFLAELRKAPVNVNCSIAVCORIQ 1019
Db 975 VELNQEAVMQVDESHVHPQNPFLSCSSSEKIPASDFLAHQKPNVDCSIAGCLFRCDV 1034
Qy 1020 PFGIOEENATLKNLSFQVYIKTSNNHLLIVSTABILNDSVFTLLPQGAFAVRSOTE 1079
Db 1035 PSFVQSELDFTLKNLSFGVVRQILQKQVSVVSVABITDTSVYSQLPQGAFAVRSOTE 1094
Qy 1080 TKVEPFPVPLPLIVGSSVGGILLALITAAALYKLGFKRQYKDMSE 1128
Db 1095 TVLEKYKHNPPTPLIVGSSIGGLLLALITAVLYKVGFFKRQYKEMME 1143

RESULT 5
Q9QXHA PRELIMINARY; PRT; 1169 AA.
AC Q9QXHA;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGA9.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Teuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
EMBL; AF211864; AAF23492.1; -.
```


Db 21 NLDVEPIVFPREDAASFGQTVVQFGSRLVVGAPLEAVAVNQTGRLYDCAPATGMCQPIV 80
QY 62 LOVPVAVNMSLGLSLAATTSPPQACGPTVHQTCTSENTYKGLCFPGSNLRQOPKP 121
Db 81 LRSPLAVNMSLGLSLVATNNAQLACGTAQACVKNYAKGSCLLIGSSL-QPIQAV 139
QY 122 PEALRGCPQSDIAFLVWGSGSIIIPHDPRRAKEFTSTWEOQKSKTILFSLMOYSBEPR 181
Db 140 PASMPCPEQEMDIAFLIDGSGINQDFAQMDKFVAKLMGEPASTSTFLSLMOYSNLIK 199
QY 182 IHTEKEFQNNPNSRLKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILIT 241
Db 200 THFTIPEFNILQSDLDIVOLQGLTYATGIRVWELFHNSKNGSRSAKILLVIT 259
QY 242 DGSKFGDPLGYEDVIEADREGVIRYVIGVGDAPRSEKSRQELNVTASKPPRDHVFQINN 301
Db 260 DGOKYRDPLEYSDVIPAADKAGIRVAIGVGDAPQBPATKELNLTGSAAPPQDHVFKVGN 319
QY 302 FEALKTIQOLREKIFAIETGOTGSSSPHEHNSQEGFSAATNSGPIILSTVGSYDMWGS 361
Db 320 FAALRSIQOLQEKIFAIETGOTGSSSPHEHNSQEGFSAATNSGPIILSTVGSYDMWGS 379
QY 362 VFYLTSEKSTFTNMTVRVSDMDNDAYLGAAAAIILNRRVQSLVGLGAPRYQHIGLVAMPRQ 421
Db 380 AFLYPENRTPPTNNSQENVMDRSDYLGYSTAVAFKGVHSLILGAPRQHTGKVVIFQ 439
QY 422 NTGWNESNANVKTQIGAFPGASLCSVDVDSNGSTDLVILGAPHYEYQTRGGQSVCPPLP 481
Db 440 EAEHWPKEVRGTQGSVFGASLCSVDVDRDSTDLVILGAPHYEYQTRGGQSVFPFP 499
QY 482 RQORARWQCDVILYCGQGPWGFAGALTVLGVHNGDKLTDVAIGAPGEDNRGAVYLP 541
Db 500 -GVYRGWQCEATLGGQGWGFGVALTVLGVNGDNLADVAIGAPGEESRGAVYIFH 558
QY 542 GTSGSGISPSHSRIAGSLSPLOVFGOSLSGGQDITMDGLVDLVGAGQGHVLLRSOP 601
Db 559 GASRLAIMPSFQSVRTGSLSLQVFGOSLSGGQDITMDGLVDLVGAGQGHVLLRSOP 618
QY 602 VLKVAIMEFNPREVARNVFECDNVVVKGEAGEVRLVOKSTRDRRLREGIOQSVVTV 661
Db 619 LLKVELSIRFAPNEVAXVQWERTPTVLEAGEATVCLTVHKGSPDLL--GNVQSVRY 676
QY 662 DLALDGRPHSRVAFNETKNSRTOVLGTOTCETLKLQLPNCIEDPVSPVILRLNPS 721
Db 677 DLALDGRILSRAIFDETNCITLGRKTLGLGPHCETVKLLPDCVEDAVSPIILRLNPS 736
QY 722 LACTPLSATGNLBPVLAEDAQRLPTALPFEKNCNDNICODDLSTPFSMSLDCILVVG 781
Db 737 LVDRDASP--NLHPVLAVSGQDHITASLPFEKNCQKELLCEGLDGFISFPGQLVVLWG 795
QY 782 PREFNVTYVRNDGDSYRTQVTFPPFLDLSYKVSQTLQNRQSQRSWRLACASSTEYS 841
Db 796 SPBLTVTVTWNEGDSYGLVAKFYYPAGLSYERVGTQ-QPHQYPLRLACAEPAQED 854
QY 842 GALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETPQ 901
Db 855 --LESSSCSINHPFIPREGATFTMITFDVSYKAFGLGRLLLRKAKSENKNDPTNKTAFQ 912
QY 902 LELPVKAVVWVTVSHGVSTKYLNFTASENTSR-VNQHVOVSNLQORSLPISLVLPV 960
Db 913 LELPVKTVYVTLISROEDSTNHNFSHSGRQEAHRYVNNLSPLKLVAVNFWFV 972
QY 961 RLNQTVWRPQVTPENLSS--TCHYKRLPSHSDFLAKRLKAPVNVNCSIAVCRIQCDI 1019
Db 973 LLNGVAVWD---VTLSSPAQGVSCVSKMPQPPDFLTQIRRSVLDCSIADECLHFRCDI 1029
QY 1020 PFGIGEEFNATLKNLSFDWYIKTSHNLLIIVSTAEILPNDSVFTLLPQCGAFVRSQTE 1079
Db 1030 PSLDIQDELDFILRGNLSPGWSQTLQEKVLLVSEAEIIFDTSVYSQLEQGEAFLEAQVE 1089
QY 1080 TKYSPFVFNPLPLIVGSSVGGILLALITALYKLGFFKQVQKDM 1126
Db 1090 TTLEEVVYVEPIELVAGSSVGGILLALITVVLVYKLGFFKQVQKEML 1136

RESULT 7

Q9WTV4 PRELIMINARY, PRT, 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZ/2J; TISSUE=Spleen;
EA Ma R. Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VMPADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.1%; Score 1534.5; DB 11; Length 1161;

Best Local Similarity 34.0%; Pred. No. 2.3e-104; Indels 101; Gaps 37;

Matches 399; Conservative 217; Mismatches 456; VVVGAPGE--GDNTGGYHCTSEFCQP 79

QY 1 FNLDTENATPQENA-RGFQSVVQLQSRVVGAPQEIIVAAQRSGLYQCDYSTGSCRP 59
Db 24 YNLDTRPTQSLAQRHFGYVLIQIEDG-VVVGAPGE--GDNTGGYHCTSEFCQP 79

QY 60 IRLQVPVAVNMSLGLSLAATTSPPQACGPTVHQTCTSENTYKGLCFPGSNLRQOPQ 119

Db 80 VSLH-GSNHTSKYLGMTLATDAKGSLLACDPGLSRTCDQNTYLSGLCYLPQSLGPM 138

QY 120 KPFEALRGCPQSDIAFLVWGSGSIIIPHDPRRAKEFTSTWEOQKSKTILFSLMOYSEE 179

Db 139 QNRPAVQECMKGVLDVFLFDGSQLDRKDFEKILEFMKDVNRKLSNTSYQFAAVQFSTD 198

QY 180 FRIHTFKEP-QNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILI 238

Db 199 CRTEFTFLDYKQNKQPDVLLGSVQPMFLNTFRATNYVVAHVPEKESGARPDATKVLV 258

QY 239 LITDGEKFGDPLGYEDVIEADREG-----VIRYVIGVGDAPRSEKSRQELNVTASKP 291

Db 259 IITDG-----EASDKGNISAAHDITRYIIIGIKHFVSQKQTLHIFASEP 304

QY 292 PRDHVQIINNFEALKTIONQREKIFALEGTQTSSTSSFEHMSQEGFSAATNSGPIILS 351

Db 305 VEEFVKILDTPEKLDLTDLQRIYAIETGTRQDLTSFNMLSSSGISADLSKGHAYVG 364

QY 352 TVGSDVWAGGVF-LYTSMEKSTFTNMTVRVSDMDNDAYLGAAAA-IILNRRVQSLVGLGAPR 409

Db 365 AVGAKNAGGLDLREDLOGATFVQEBPLTSDVRGGYLYTVVWMTSSRSLPLAAGAPR 424

QY 410 YOHIGLVAMFR-QNTGWNESNANVKTQIGAFPGASLCSVDVDSNGSTDLVILGAPHY 467

425 YQHVGVQLLFOAPEAGGRWNQTKITGTQIGSYFGBELCSVDLDQDGEALLIGAPLFF 494

468 EOTRGQGVSCVPLPRGORARWQCDVLYGEOGQPGWGRFGAALTVLGVDVNGDKLTDVAIGA 527

485 GEORGGRVPTY---QRRQSIFEMVSELQDGPYPLGRFGAAITALTIDINGDRLTVDVAVGA 541

528 PGEEDNRGAVYLPHTGTSGGISPSHSHQRTAGSKLSPQLQYFGQSLGGQDLTMDGLVOLT 587

542 PLTEE--QGAVIITFNGKPG-GLSPQPSQRIQGAOVFFGIRWFGRSIHGVKDLGGRLDVV 598

588 VGAQGHVLLLRSGPVLKRVKXAIMEFNPREVARNVPECDNQVVKGEZAG-EVRVCLHVRQKST 646

599 VGPEGRVVLSSRPVDDVVTLSFSEPEEIPVHEVECSYAREQKHGVKLKACFRKPLT 658

647 RDRLRGQIQSVVTDALDGRPHGRAVFNKTNSTRTQTVGLTQTCTETLKLQLPNC 706

659 PQ--FOGRLIANLYTLQLDGHRMRGRLEPDSGSHELSGNTSITP-DKSLDFHPHPFPC 715

707 IEDPVSPIVLNFSIV---GTPLSAFGN-LRPVLAEDAQRLTALTALPPPEKNGCNDNICO 762

716 IQDLISPINVLSNFSLLBEEGTPRQKGRAMQILRPSIHTV-TKEIPPEKNGEDKKCE 774

763 DDLISITFSMSLDCLVVGSP-----REFNVTVTVRNGEDSYRTQVTFPPFDLDSVRKV 816

775 ANLTLSPPARS-----GPLRLWSSASLAVEMTLSNSGEDAYWVRDLDFPRGLSPRKV 827

817 STLQNRQSQRSLACESASTVEUGAL-KSTSCSINHPIFFPENSEVTFNITFDVDSKAS 875

828 EMLQ---PHSRMPVSCUEL--TEGSSLLITLKNVSSPIFKAGQEVSLQVMNTLINS 882

876 LGNKLLIKAMVTSEN-NMPRTNKTETFOLELPYKIVAVMVVTSHGVSKYLNFTASRNTSR 934

883 WEDPFVNLGTVHCENENSSIQEDNSAATHIPVLPVNVILTKQENSTLYISFTPKGPKTQ 942

935 VMOHQVQV---SNLQBSLPISLVFLVPVRLNQTVIHDRPQ--VTPSENLS--TCHT 985

943 QVQHVTQVRIQPSAYDHNMP-TLEALVGVP-----WPHSEDPITYTWSVQTDPLVTCHS 995

986 KE-RLPSHSDFLAELRKAPVNVGNSIAVCQRIQCDIPFGIQEERFNATLKGNSLFDWYIKT 1044

996 EDLKRPSSE--AEQCLPGV-----QRCPIVE--RREILIQVTGVLSKEIXA 1041

1045 SRNHLLIVTAELIENDSVFTLLPGQAGFVRSQTETKVFEPVNPPLPIVGSVVGGLL 1104

1042 S-STLSCSLSVSFNSKHFHLYGSKA-SEAQVLVKVDLIHEKEMLVVYLSGIGGLVL 1099

1105 LALITAALYKLGPFRQYKDMW-SEGPPGGAEP 1136

1100 LFLIFLALYKVGFFKMLKGRKADGGVPNGSP 1132

RESULT 8

Q9R200 PRELIMINARY; PRT; 1160 AA.

AC Q9R200;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Integrin alpha L.

GN ITGAL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TissUE=Spleen;

RA Ma R. Z., Teuscher C.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF065901; AAD25884.1; -.

DR HSSP; P20701; ILFA.

DR MGD; MGI:96606; Itgal.

DR GO; GO:0008305; C:integrin complex; IEA.


```
Db 775 ANLTSSPARS-----GFLRLMSASLAVETWLTNSGDAYVWRDLDPFRLGSLFRKV 827
Qy 817 STIQNQSQRWRLACESASSSTVSAL--KSTSCSINHGIFFENSHVYTNITFDVDSKAS 875
Db 828 EMLQ-----PHSRMPVSCBEL--TEGSLTTLTKCNVSSPIFKAGQSVSVQVMFNILLNS 882
Qy 876 LGNKLKLLKANVTSEN-NMPTNKTEPQLPVLKVIAYVMVVTSHGYSKYNFTASENISR 934
Db 883 WEDFVELNGVHCENENSSLOEDNSAATHIPVLVPVNIUTKQENSTLYISFTPKGPKTQ 942
Qy 935 VMOHQOVNGLGORSIPISLVELVRLNQTVWDRPQ-----VTPSENLS-----TCHTK 986
Db 943 QVQHVTVQ-----RIQFSAYDRNMPT-LEALGVPRPHSEDLITYTWSVOTDPLVTCHE 996
Qy 987 E-RLPSSDFLAELRKAAPVNVCSIAVCQRIQCDIPFGIQQEFNAFLKGNLKFDFWIKTS 1045
Db 997 DLKRPSS-----SEAPCLPGV--QFRCPIVE--RWELLIQVTGVELSKERAS 1041
Qy 1046 HNHLLIVSTABILFNDSVFTLLPGQGFVRSQTEKVEFFEPNPLPLVIGSVVGLLL 1105
Db 1042 -STLSLCSLSVSFNSKHFHYGSKA-SEAQVLKVDLIHEKEMLHVYVLSGIQGLVLL 1099
Qy 1106 ALITAAALYKLGPPKROYKDMW-SEGGPPGAEP 1136
Db 1100 FLIFLALYKVGFKMLKEMKMEADGGVPNGSP 1131

RESULT 9
ID Q98TF1 PRELIMINARY, PRT; 1196 AA.
AC Q98TF1;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CD11-1.
GN CIA1.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB048536; BAB39134.1; -.
DR HSSP; P20701; 1LFA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PFO1839; FG-GAP; 3.
DR Pfam; PFO0357; Integrin_A; 1.
DR Pfam; PFO00392; wwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SMC0191; Int_alpha; 5.
DR SMART; SMC0327; WWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50234; WWA; 1.
SQ SEQUENCE 1196 AA, 132477 MW, 9369C807E7DCA53B CRC64;

Query Match 24.0%; Score 1409; DB 13; Length 1196;
Best Local Similarity 31.5%; Pred. No. 5e-95;
Matches 375; Conservative 224; Mismatches 447; Indels 144; Gaps 39;

Qy 1 ENLDTENAMTFQENARG-FQGSVVQLQ-GSR--VVVGAQEIVAANQRGSLQCDYSTGS 56
```

```
Db 32 FNIDTEPLRPMGAPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSGEMYSCTADLQS 88
Qy 57 CEPRLQVP-----VEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLRGS 112
Db 89 CQ--RLQRPGESVRFGMSAAVSSAALTS-----CSYPFHEDCDGNSYLNGVCYQFSS 140
Qy 113 NLROQPOKFPBALRCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMBOLKSKSLTFS 172
Db 141 SL-QAVGNFTAAQBCSKREVLNVLFDGSSSMKTVDFEMKNFKIKDLMKLSNSSIKFA 199
Qy 173 LMOYSEPRTHFTKEPKNPNPSLIKPIITOLLGRHTATGIRKVVRELEN-ITNGARK 231
Db 200 AVQSTIDVTVDFNDYQSGSAEKLAKB--THMKSINTHKADIDYILKLNLSLMSGADS 258
Qy 232 NAFKILILITDGEKFGDPLGYED--VIPEADREGVIRYVIGVGDAFASHKRSQELNIVAS 289
Db 259 KAQKALVIITD---GDPDNDNDYNLKKCKDEQNILAYIIGV---KVDLIELTQLAS 309
Qy 290 KPRDHPVPOINNFAALXTIOMQLSEKIPATBGTOTGSSSSSEHEMSQEGRSAAITSNGL 349
Db 310 EPEKNFTYIKDYSGLGKLLNLOKTYNIBGSYDAQGRDRQKELSGSGFSVYQEBESVI 369
Qy 350 LSTVGSYDWAQGVFLYTSKEKSTETINMTRVDSDMN-DAYLGYAAAIILNRVQSIVLGP 408
Db 370 VGSVGSNDWRGALYEVTG--SGSDPKETIIDPAVNKDSYMGYSTVLGHRGVSLFSGAP 428
Qy 409 RYOHIGLVAMPRONTGMWESNANVKGTQIGAYPGASLCSDVDVDSNGSTDLVLIGAPHYE 468
Db 429 RAETGLVTLFTKNTQNTVTVTSNINEQIGSYFGASLSLLDSDGSDGDFLLVGLAPFYQ 488
Qy 469 QTRGQSVSVCPLPRGORARWQCDVALYGEQ-----GQPWGRFGAALTVLGVDWNGDKLTDV 523
Db 489 SQ-----PRTGRLVYVLSLSEQYFQKTLNYSQSTTGTPAASVASLKDNGDGLSDV 540
Qy 524 AIGAPGEDNREGAVYLPHTGTSGSGISPSHS--QRIAGSKLSPLQYFGOSLGGQDLNDG 582
Db 541 AVGAPLE--NEGVMVYVILGDRTHGINBELTPQIRISVQSIVLPGAQFGVSLTGQMDMNDN 598
Qy 583 LVDLTVGAGHVLLRSQPLVKALMEFNPREVARNVFECDNVQVKGKAGEVRVCLHV 642
Db 599 LTDIVIGAQGGIVLLKARPVMSVAQSUSFKREISLNTFECPGS--NAFNAFLTSCFTV 656
Qy 643 OKSTRDLREGQIQSV--VTYDLALDSCRPHSRVAFNETKNSTR--ROTQVLGLTQTCT 698
Db 657 TERISS---TGSLEKLNVLNLDVVDVVRGMSRGFFDQSSVSSRTLQOSVLLDSGSSCFN 713
Qy 699 LKLQLPNCIEDPVPSPIVLRNLSLVGTPLSAFGLNLRPVLAEDAQRLLFTALPPEKQCGND 758
Db 714 FSIFMLRCVADTVSPLKIRNFS--QTEMLS-GNSVAVLDVHSRTEENVEVFFQRCNCSN 770
Qy 759 NIQDDLSITSPFMSLDCLVVGGPREFNVTVTVRNDGBDSYRTQVTPFPFLDLGYRYKYST 818
Db 771 NSCVADLKLNFSPTN-WTLVVNQAHFTVQVSLANPGDDSYNTSIVLHYPGISLSKXPA 829
Qy 819 LQONRSQRWRLACESASSTVSALKSTSCSINHPIFPENSEVTFNITPDV---DSKAS 875
Db 830 IKPSRTR-----SSCGDRSDGATNRTTCSIDLVPYRSGTTTQFLGTFPVMKWDNDS 881
Qy 876 LGNKLKLLKANVTSENMPRTNKTTFQLELPVKYAVVMVVTSHGV-STKYLNFASENISR 934
Db 882 NMEIMITANDNDNGM---SDTEVRSSVFPQPAVDLAISLVAEDSVTYMNFSLDRGPK 938
Qy 935 VMOHQOVNGLGORSIPISLVELVRLNQTVWDRPQ-----VTPSENLS-----LN 963
Db 939 PLNITYKVNGSRKOLPVSVTLTLFCQTPHVILTPHFTPSMHEVHSFTSSYHQIMCLLN 998
Qy 964 QTVIWRPQVTFSENLSTCHTKERLPSSHDFLAELRKAAPVNVCSIAVCORIQ-----C 1017
Db 999 KELLFPSELGAQVQRTGWSLREVEC-SQF--DLNKSVAHNLITADALQNKEYES 1055
Qy 1018 DIPFFGIOBE--FNATLKGNLSFDWYIKTSHNLLIVSTAILFNDSVFTLLPGQGFVR 1075
```

```

Db 1056 KTSYFVFRKDNVFSISAEINNTSYNQTSS-----SELKYNPH-----R 1094
Qy 1076 SOTETKVEFEVFNPLPIIV-GSSVGGLLLLALITAAALYKLGFPKROYKD 1124
Db 1095 SQTQVKVE-FVVPSPSLMLIVCTGAVGGPFPIILFLLLKCGFFKRNRPD 1143

RESULT 10
Q96HB1
ID Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AAH08777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWEA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6PF2546E8C632P9 CRC64;

Query Match 23.1%; Score 1359.5; DB 4; Length 1086;
Best Local Similarity 32.3%; Pred. No. 2e-91;
Matches 378; Conservative 185; Mismatches 431; Indels 175; Gaps 37;

Qy 1 ENLDTENAMTEQ--ENARGFQSVVQLQSGRVVVGAPQELVAORGSIXOCYDSTGSC 58
Db 26 YNLVDRGARSPPPRAGRHFGTRVLQV-GNGVINGAPGE---GNSGSLYQCSGTHCL 81
Qy 59 PIRLQVPVEAVNMSGLSLAATSPQQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQP 118
Db 82 PVTLR-GSNYTSKYLGMTLA--TDP----- 103
Qy 119 QKPEALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEPISTVWEQKSKTLFSLMOYSE 178
Db 104 -----TDGS-----ILFAAVQSF 117
Qy 179 EFRIHFTKEFQNNPNRSLIKPIQLLGRTHATGIRKWKRELPIITNGARKAPKILI 238
Db 118 SYKTEFSDYVYKRDPAKLLKHVKHMLLTTFGAINVATEVEFRELGARPDATKILI 177
Qy 239 LITDGEKGDPLGYSDVLPADREGVIRYVIGVGDAFRSEKSRQELNVTASKPRDHFVQ 298
Db 178 IITDGE--ATDSGNTDAKD-----IIRYIIGHKHFTKESQETLHKFASKPASEFVKI 230
Qy 299 INNFALKTIONOLREKIFALEGTGTGSSSEFHEMSQEGFSAATSGNPLISTVGSYDW 358
Db 231 LDTFELKDLFELQKLYIVTEGTSKQDLTSFNWELSSGISADLSRGHVVAVGAKDW 290
Qy 359 AGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAAA-IILNRVOSLVLGAPRQHIGLV 416
Db 291 AGGFGLDKADLQDDTFIGNPELTPVVRAGYLGYVTWLPSPKQKTSLLASGAPRYQHMGRV 350

```

```

RESULT 11
Q98TF0
ID Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CD11-2.
DE CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";

```

```

Qy 417 AMFR--QMTGWEGNANVKGTOIGAYFGASLCSDVDNSGSTDVLVIGAPHYETRGQ 474
Db 351 LFFORPQGGHWSQVTHGTQISYFGELCGVDVDGGETELLIGAPLYGQRQGR 410
Qy 475 VSVCLPRGQBARWQCDAY--LYGEQGPWRFGAALTVLGDVNGDKLTDVAIGAREED 532
Db 411 VFYI-----QRRQLGFEVSELOQDPGYPLGRPEAITALTIDNGDGLVDVAVGAPLEE- 464
Qy 533 NRGAVYLFHGTSGSIGSPSQRISAGSKSLPRLQYFGQSLSGGQDLTMDGLVDLTVGAQG 592
Db 465 -QGAVYIFNGRHG-GLSPQPSQRLEGTVLGGIOMFGRSHGVKDLBESGLADVAGAES 522
Qy 593 HVLRLSRQVLRVKAIMEFNPVARNVFNCDQV--KGEAGEVVRVCLHVWKSTRDLRL 651
Db 523 QMVLSSRPVDMVTLMSFSPAEIPVHEVECSYSTSNKMEGVNITICFQI-KSLIPQF- 580
Qy 652 EGOQSUVTYDLALDSGRPHSRVFNSTKSTQTOVLGTOCTETLKLQLPNCIEBPV 711
Db 581 QGRIVANLTYTLQDGHTRRRGLFPGGRHRLRNIAVT-TSMSCDPSFPFPVCVDLI 639
Qy 712 SPIVRLNPSL---VGTPLSAPGN-----LRPVLARDAORLFTALPFPFNKCGNDNICQ 762
Db 640 SPINVSINFLWREGEPTRDORAGKDIPPLRPSLHSETWEI-----PFEKNCGEDKKCE 694
Qy 763 DLSITSPMSLDCLVGGPREFNVTYVRNDGDSYRTQVTFPFDLDSYRKYSTLQNQ 822
Db 595 ANLRVSFSPAKSRALRUTAFASLSVELSLNLEDAYVQDLHPPLGLSPKRVEML--- 751
Qy 823 RSQSRWLACES--ASSTEVSGALKSTCSINHPFPENSEVTFNITFDVDSKASLGKL 880
Db 752 KPHSQIPVSCBELPEBSRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSSWGSV 808
Qy 881 LKXANTVSENN---MPTNKTPEQLELPVKYAVVYVTVSHGVSTKYLNTASENTSRVM 936
Db 809 ELHANVTCCNEDSLLDENSATTI---IPIDYINILIQOEDSTLVSVFTKPGKIHQV 865
Qy 937 QHOYQV---SNLQORSIP-LSLVLVPLVRLMOTVINDRPQVTFSENLSSTCHTK--BRLP 990
Db 866 KMYQVRIQPSIHDNIPTLEAVGVGPQPPSEGIPTQWSVQMEPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVNCSLAVCQRIQCDIPFGIOEBFNATLKMISFMYIKTSHNH 1048
Db 924 DAAPCPALGALFRCPVV-----PRQILVQVIGTLELVGEIEAS--SM 964
Qy 1049 LLIVSTAEILFNDSPVFTLLPGQAFVRSQETKVEPPEVFNPLPLIVGSSVGGILLALI 1108
Db 965 FSLCSSLSISFNSSKPHLYGSNASL-AQVVMKVDVYVYKQMLYLYVLSGIGILLILLI 1023
Qy 1109 TAALYKLGFPKROYKDMMSSEG-GPPGABP 1136
Db 1024 FIVLYKVGFTFRNLKERMEAGRGVNGIP 1052

```

RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB04537; BAB39135.1; ..
DR	HSSP; P20701; INF.
DR	GO; GO:0008305; C:integrin complex; IEA.
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR	InterPro; IPR001969; Asparticase AS.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP; 3.
DR	Pfam; PF00692; vwa; 1.
DR	PRINTS; PR01185; INTEGRINA.
DR	PRINTS; PR00453; VWFADOMAIN.
DR	SMART; SM00191; Int_alpha; 5.
DR	SMART; SM00327; vwa; 1.
DR	PROSITE; PS00141; ASP_PROTEASE; 1.
DR	PROSITE; PS0234; VWFA; 1.
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDCTCA8B6B1C59 CRC64;
Query Match 23.1%; Score 1358.5; DB 13; Length 1187;	
Best Local Similarity 31.6%; Pred. No. 2.8e-91;	
Matches 377; Conservative 206; Mismatches 452; Indels 157; Gaps 43;	
QY	1 FNLDENAMTFQENARG-FGQSVVQLQ-GSR--VWVGAQOEIVAAQNRGLYQCDYSTGS 56
DB	32 FNIDTEPLRFNGTPEDFGVSVVQTEFGNKKQLIIVGAPLE---GNSAGEMYCTADLQS 88
QY	57 CEIRLOVP-----VEAVNSLGLSLAATTFPPQLLACGPTVHTCSENTVKGCLFLGS 112
DB	89 CK--RLQRPQSESVRFPGNSAAVSSAALTS-----CSPYFAHECGNSYLNGVCYQFNS 140
QY	113 NLRQPKPPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKBFISTVMEQLKSKTLFS 172
DB	141 SL-QAVSNFTAAVQECSEKREVNLFVFDGSSMKAVDFQNMKNFKIDVMMKLSNNSIKPA 199
QY	173 LMQVSEFRHFTPKFQNNPNPRSLIKPTOLLGRTHATGTRKVVRELFN-ITNGARK 231
DB	200 AVQFSTEIRTVDFDNYQNGSAEELMKR-RHMKSLNTYKAINYLKYNLNSVSGADP 258
QY	232 NAFKILLITDGERFGDPLGYED--VPEADREGVIRYVIGVDGAFRSEKSRQELNLTAS 289
DB	259 NAQVALVIITD-----GPSNDNDYVILNICEQNILYIIVGV-----KYDLITLTQLAA 309
QY	290 KPPRDHVQINNFPALKTIQNLREKIFAEGTGTGSSSFEHMSQEGFSAITSNGPL 349
DB	310 EPKLNNTFYIQEYNGLGLLDNLQKIYNIEGSKAEHGRDRQKELSQSGFSVWYQEESVI 369
QY	350 LSTVGSYDWAGGFLYT---SKERSTFINTRVDSMDNDVAGYAAAILNRNRVQSILVG 406
DB	370 VGSVGSNDWRGALYEVMGSGSKFQTEITDPAVN---KDSYMGISTVLGKRGHGVSLPSG 426
QY	407 APYQHIGLVAMFRQNTGMBSNANVKGQTQIGYFGASLCSDVDVDSNGSTDVLVIGAPHY 466
DB	427 APRAHTGLVTLPTKNBSTWTVMENINGEIGSYFGASLSLLDSDSDSDFLVIGAPLF 486
QY	467 YE-QTRG-GQVSVCLPFGQPARWQCDVLYGEO--GQPWGRFGAALTVLGDVNGDKLTD 522
DB	487 YCSQPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDNLGDLSD 535
QY	523 VAIGAPGEDNRGAVLPHGTSGSGISPSHS-QRIAGSKLPRLYQFCQSLSGQCDLTMD 581
DB	536 VAVGAPLE--NEGVVYIIGDGTGHINFEHAPQRIPARSVLPGLQFGVSLSGQMDMND 593
QY	582 GLVLDUTGAQGHVLLRQPVLRVKAIMEFPREVARNVFECDNQVVKGEAGEVRVCLH 641
DB	594 NLPDIVIGTGGVILLNARPNVMSYSAQLSFNPMRISLNYFPCGS--NAPNAPNLTSCFT 651
QY	642 VOKSFDRLRQGIQSV--VTVDLALDSGRPHSRAVFNETKNSTUR--RQTVGLGLTQCE 697
DB	652 VTERISS---TGSLEKKLVNSLNLNDVVRGKSRGFFDPDSDSSRTLQOSVILLDSGSCS 708
QY	698 TLKQLPNCIEDPSPVIVLRNFS---LVGTPLSAFCGNLRPVLAEQAQRLLFTALFPPEK 753
DB	709 NFSIFMLECVADTVSPLKIRNFSQTKLSCNSLAVL-DIQSRTEYYEVL-----PQR 761
QY	754 NCGNDNICQDDLSITFSGMSLDCLVWGGPREFNTVTVRNDGDSYRQTQVTFPPDLDSY 813
DB	762 NC-NSNSCVADKLNFSTN-DTLVVENQAHTVTVLSLANFGDDSYNTSIVLHYPEGUSL 819
QY	814 RKVSTLQNRORSRWLACBSASSTEVSGALKSTSCSINHPIFFPENSEVTNITPDV--- 870
DB	820 SKFAIKPSRTR-----SSCGDRDSGATNRTCSINLFPYRSGTTTQFLGTFPRVTKW 871
QY	871 DSKASLGNKLLKANVTSENMPRTNKTFOLELPVKYAVVWVTVSHGV-STKYLNFAS 929
DB	872 DYDWSDRMEMTITANSNNGNM--SDMSVRSIPVQFAVELAISLVAEDSVTYLNFSL 928
QY	930 ENTSRVMOHOVSNLGORSIPISLAVPLVPVRLAQTVIWDPRQVTFSENLSTCHTKERL 989
DB	939 DRGPKPLNIYKVNHLGKLPVSVTLSPCO-----TTHVLTLPENFSMQ 974
QY	990 PSHSDFLAELRKAPVVC-----SIAVCORIQCDIPFFGIOEE 1027
DB	975 EVHHSFISSYHQ--IIMCLLNKHLFFSPSLSAVQTKTGRSLWVC-----VSSISTGBI 1026
QY	1028 FNATLKGNL-----SFDWIKTSHNHLIVSTAELFNDSVFTLLPQ-OCA 1072
DB	1027 FRSSV--NLMAEAVLQNVKEYESKYSFY-EFRDHFVFNIS-AELAFNTSRYNQSGTLKYN 1082
QY	1073 FVRSQTEVKPEFPEPNPLIVGSSVGGLLLLALITAALYKLGFYKQYKD 1124
DB	1083 PHRSQTEVKPEFVPPSRMLIVCTGAVGSGFPPLIILILLKCGFFKMRPD 1134
RESULT 12	
Q8HZVO PRELIMINARY; PRT; 927 AA.	
AC	Q8HZVO;
DT	01-WAR-2003 (TrEMBLrel. 23, Created)
DT	01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Lymphocyte function-associated antigen 1 (Fragment).
OS	Bos taurus (Bovine).
OC	Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT	"Sequence of the alpha subunit of bovine lymphocyte function-
RT	associated antigen 1."
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; A2440778; AAN63636.1; ..
DR	PIR; A32039; A32039.
DR	GO; GO:0008305; C:integrin complex; IEA.
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP; 3.
DR	Pfam; PF00092; vwa; 1.
DR	PRINTS; PR01185; INTEGRINA.
DR	SMART; SM00191; Int_alpha; 4.
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE; PS0234; VWFA; 1.
FT	NON_TER 1 1
FT	NON_TER 927 927
SQ	SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;
Query Match 21.7%; Score 1278; DB 6; Length 927;	
Best Local Similarity 34.9%; Pred. No. 1.7e-85;	
Matches 341; Conservative 166; Mismatches 385; Indels 86; Gaps 28;	

RESULT 13
Q88340
ID Q88340

489 - EDAFVRIRIEGQMSYFGSVLCFVDIDMDGTTDFLLVAAPFYHIRGEGRVYVTVQPE 547
483 GQARWQCDVLYGEGQFQWGRFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
548 - QDASFSLAHTLSSHPCLNTSRFGFMAAVGDINQDKFTDVAIGAPLEFGAGDGASYS 606
537 VYLFPHGTSGGSIPISSHQRTAGSKLPLOYFGQSLSGGDLTMDGLVLTWCAQGHVLL 596
607 VYIYNGHSG-GLYDPSQIRASSVAGSLHYFGKSVSGGLDFNGDGLADITVGRSDAVV 665
597 LRQOPVLKVAIMEFNPVARNVFCNDQVVGKEAGEVRVCLHVQKS-----TRDLREG 653
666 LRSPVVDLTVSMITFP-----DALPWFVIGKM--DVNLCFEVDSSVASEPGLREM 715
654 QIOSVVTYDIALDSGRPHSRVAVNETKNTROTQVGLTQTC----- 696
716 FLNFTVDVY-----TKQORLQCEDSSGCGCLRWKNGSFCCEHFWII 760
697 ETLLKQLPNCIEDPVPSPVLRNLFSLVGTPLSAFQNLK-----PVLAEQAORLETFALF--P 750
761 STEEL-----CEEDCFSNITIKYVE-----PQISGGERDYENFTL--DHYKEPSAIFQLP 809
751 FEKNCGNDNTQQDLSITTFMSLDCLVUGPREFNTVTVRNDGEDSYTQTVPFFPLD 810
810 YEKDCKNVFCIABIQLTIN--ISQQLVVGVTKEVTMNIISLTNSGSDSYMTNMAIYPRN 868
811 LSYRKVSTLQNRQSRWRLACASASTEVSGALKSTSCSINHPIPFENSEVENTEDV 870
869 LQFKK-----QXPSPDQCDPKV---ASVLWNCKLGHFIL--KRSVNVSVTWQL 918
871 DSKASLGNKLLKANVTSENNMPNTKTEFQLELPKYAYMVVTVSHGVTKYLNFTASE 930
919 EESVFNRTADITVITISNEKSLARETR--SLOPRHAFIAPVLSR--PSVMYMN--TSQ 971
931 NTSRVMOHQVQVNLGORSPLISLAVFLVPLVRLNQTIVDRPQVTFSENLSST-----CHT 985
972 SPSPDKHFFNVHGENLFGAVFQIQICVPIKLOF-----QIVRVKNLTKTQDHECTQ 1025
986 KERLPESHDSFLAEIRKAPVNVCSIAVCQRIQCDIPFEGIOEBEFNATLKGNSLDWYIKTS 1045
1026 SQEPACGSDPQVHKWHSVVCAI-----TSKNVTVAAEISVG 1065
1046 HNHLIVSTA-----BILFNDSVFTLLPGGAFVRSQTKKEPF-----EVENPLPLIV 1095
1066 HTKQLLRDVSELPIGSEISFNKSLYGLNAB-----NHRKTIVIFLKEETRSPLII 1119
1096 GSSVGGLLLLALITAAALYKLGFFKQVQKMMSE 1128
1120 GSSIGGLLVVVIILALFKCGFFKRYQQLNLE 1152

RESULT 14
O88341
ID O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98056920; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.,"
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.

GO; GO:0008305; C:integrin complex; IEA.
GO; GO:0004895; P:cell adhesion receptor activity; IEA.
GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
FT NON_TER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;

Query Match 19.1%; Score 1123; DB 11; Length 1167;
Best Local Similarity 27.7%; Pred. No. 8e-74; Indels 244; Gaps 37;
Matches 334; Conservative 207; Mismatches 421;

QY 47 LYCCDYSTGSCBPIRLQVPEAVNMVSLGLSLAATTPSPOLLACGPTVH----- 94
DB 67 LIQC-----ATSIIDEICQCPVEHILMPKRGYQVT 98

QY 95 -----QTCSENVTYKGLCFPGNLRQOPKPPLEALG----- 127
DB 99 LVNRHNGVLVCTIQVQSRKPRSLNSELTGACSLTLPNLDLQAAQYFSDLEGLDLGASVNS 158

QY 128 -----CPOED-----SDIAFLVDGSGSIIPHDFFRAKPI 157
DB 159 GDIYSGKSGSTGEETYSARRLRQARVEDEDEAGTEIAVLVDGSGSIEPSDFQAKDFI 218

QY 158 STVMEQL--KSKTLFLSLMOYSEERIHTPFKEFQNNPNRSLIPIITQLLGRTHATGI 215
DB 219 STWNRNFEYKEPCFNALVQYGVQIQTPEFDLDSRDINASLAKVQSVIVQKEVTKTASAM 278

QY 216 RKVRELFINITGARKNAKILILITDGEKGDPLGYEDVPEADREGVIRVIGVGDAF 275
DB 279 QHVLNDIFIPSGSRGALKVWVLTGDFIRDPNLNLTIVISSRMQGVVRFALGVGNF 338

QY 276 RSEKSRQELNTVASKPRDRHVFQINNFEALKTQNLREKIFAIZETQTCSSSSFEHMS 335
DB 339 ENNTYRELKLIASDPKAAHTFKVTNYSALDGLLSKLORIIHMEGT---VGTLYQYLA 395

QY 336 QEGFSAITSNGP-LLSTVSGVDWAGGVLY--TSKEKSTFINNTRVDSMDNDA---YLG 390
DB 396 QTFSAQILDKGOVLGTGAFNWSGGALLTYNONGRFLNQT--AKEDFRAAQYSLGY 454

QY 391 AAIILNRNVQSLVLCAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDV 450
DB 455 SYAALHKAHGVSYVACAPRHLKRGAVFELQKEDGEBTFMRIECEQMGSYFGSVLCPVDI 514

QY 451 DSNGSTDLVLIGAPHYTYETRGQVSVCPFLPRQARWQCDVLYGECQCPNGRFGAALT 510
DB 515 NMDGITDFLLVAAPFYHIRGEGRVYVYRV--HQDAPFSLVYTLGYPLTSSRFGFAMA 573

QY 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGAVYLPFGTSGSGISPSHSQRIAGSKLSPR 564
DB 574 AVGDINQDFTDVAIGAPLEFGAGDGASVGYIYNGHSG-GLHASPSCQIRASSVALG 632

QY 565 LQYFGQSLSGQDLTMDGLVLTWCAQGHVLLRSDPVLKVAIMEFNPVARNVFCN 624
DB 633 LYIFGMSVSGGLDFSGDDLADITVGSQDVAVVLRSEPPVDLTVSMITFPDALP----- 695

QY 625 DQVVKGEAGEVERVCLHVQKS---TRDLREGQIQSWVTYDIALDSGRPHSRVAVNETKN 681
DB 686 ---MAPKQMDVELCPKVDSSAVSPSPGLRGLNSLNTVDVY-----TKQ 727

QY 682 STRRQ-----TQVLGLTQTCETLKLQLPNCIEDPVPSPVLRNIFS 721
DB 728 KORLQCADRSCCQCLMKWNGSGSLCEHFLGISTEBEL-----CBDDCFSNITIKVSYE 781

DB 932 TIPVKIEVGLIFVSVFKEHHVIAANETIPTAINTIEQIGDEVTLHYRIKGEHPMENL 991
QY 953 SLVFLVP--VRLNQTVINDRPOVTFSENLSSTCHTKE-----RLPSHSDFLAELKAP-- 1003
DB 992 TIQILYPDVTTAKNTLLY--ITTLSHSONAICKSSYPVDHLKIGSKSYVLPKIKEPTK 1048
QY 1004 --VVNCSIAVCORIOCDIPFGIOBEFNATLKGNSFDW---YIKTSHNHLLIYSTAEIL 1058
DB 1049 DTIMECDTFCASINCALAPSDI-SQVNVSLR-----VWKPTIIKASIHSLTVVVKALHR 1102
QY 1059 FNDVSFTLPGQCAFVRSQETKVEPFEVNPPL--IVGSSVGGILLIATIALYKLG 1116
DB 1103 SENSSIL---ENDHQLETMIKISKBPFGSVPLWVPLSIFAGLLILALLIPALWKAG 1159
QY 1117 FFKROYKMMSE 1128
DB 1160 FFKRPLKKOEX 1171

Search completed: June 7, 2004, 17:16:25
Job time : 43.4889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTQENARGFGQ.....PKROYKMMSEGGPGGAEQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	2	AAW65090 Human Bet
2	5862	99.8	1153	2	AAW65090 Human Bet
3	5862	99.8	1153	5	AAU80252 Human Int
4	5862	99.8	1153	5	ABG61469 Human Bet
5	5862	99.8	1153	5	AAO14428 Integrin
6	5862	99.8	1153	7	ADP25615 Binding d
7	5852	99.6	1153	2	AAW65090 Human Bet
8	3460	58.9	1163	2	AAW65091 Human Bet
9	3446	58.6	1163	2	AAW65091 Human Bet
10	3446	58.6	1163	3	AAW65091 Human Bet
11	3446	58.6	1163	5	ABG61470 Human CD1
12	3444	58.0	1163	6	ABU07406 Protein d
13	3411	58.0	1161	2	AAW65091 Human Bet
14	3411	58.0	1161	2	AAW65091 Human Bet
15	3411	58.0	1161	2	AAW65091 Human Bet
16	3411	58.0	1161	2	AAW65091 Human Bet
17	3411	58.0	1161	2	AAW65091 Human Bet
18	3411	58.0	1161	2	AAW65091 Human Bet
19	3411	58.0	1161	3	AAW65091 Human Bet
20	3411	58.0	1161	5	ABG61468 Human Bet
21	3395.5	57.8	1161	2	AAW65091 Human Bet
22	3395.5	57.8	1161	2	AAW65091 Human Bet
23	3395.5	57.8	1161	2	AAW65091 Human Bet
24	3395.5	57.8	1161	2	AAW65091 Human Bet
25	3395.5	57.8	1161	3	AAW65091 Human Bet

26	3395.5	57.8	1161	5	ABG61485	Abg61485 Human Bet
27	3234.5	55.0	1161	2	AAW78169	Aar78169 Rat alpha
28	3234.5	55.0	1161	2	AAW23062	Aaw23062 Rat beta
29	3232.5	55.0	1161	2	AAW60004	Asw60004 Rat alpha
30	3232.5	55.0	1161	2	AAW72824	Asw72824 Rat alpha
31	3232.5	55.0	1161	3	AAW07374	Abw07374 Rat alpha
32	3232.5	55.0	1161	5	ABG61483	Abg61483 Rat beta
33	3225.5	54.9	1161	2	AAW65104	Aaw65104 Rat alpha
34	3225.5	54.9	1161	2	AAW73345	Asw73345 Rat beta
35	3222	54.8	1161	2	AAW23061	Aaw23061 Mouse bet
36	3222	54.8	1161	2	AAW60003	Asw60003 Mouse alp
37	3222	54.8	1161	2	AAW65103	Aaw65103 Mouse bet
38	3222	54.8	1161	2	AAW72836	Asw72836 Mouse alp
39	3222	54.8	1161	2	AAW73347	Aaw73347 Mouse alp
40	3222	54.8	1161	3	AAW07373	Abw07373 Mouse alp
41	3222	54.8	1161	5	ABG61482	Abg61482 Mouse bet
42	3218	54.8	1161	2	AAW78168	Aar78168 Mouse alp
43	3211.5	54.7	1151	2	AAW23059	Asw23059 Rat beta
44	3211.5	54.7	1151	2	AAW60001	Asw60001 Rat alpha
45	3211.5	54.7	1151	2	AAW65101	Aaw65101 Rat beta-

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX
AC AAW65090;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11b subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11b subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC	diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,	
CC	asthma, psoriasis, lung inflammation, acute respiratory distress syndrome	
CC	and rheumatoid arthritis	
XX		
SQ	Sequence 1153 AA;	
	Query Match 99.8%; Score 5862; DB 2; Length 1153;	
	Best Local Similarity 99.3%; Pred. No. 0;	
	Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;	
QY	1 ENLDTENAMTFQENARFGQSVVQLOGSRVVVAGAPQEIIVAAANQORSLYOCYSTGSCPEI 60	QY 961 RLNQTIVWRPQVTFSENLSTCHTKERLPSHSDPLAELRKAPVNVCSIAVCQRIQCIP 1020
DB	17 ENLDTENAMTFQENARFGQSVVQLOGSRVVVAGAPQEIIVAAANQORSLYOCYSTGSCPEI 76	DB 977 RLNQTIVWRPQVTFSENLSTCHTKERLPSHSDPLAELRKAPVNVCSIAVCQRIQCIP 1036
QY	61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTPYVKGCLPFGSNLRQPOK 120	QY 1021 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQCAFVRSQTET 1080
DB	77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTPYVKGCLPFGSNLRQPOK 136	DB 1037 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQCAFVRSQTET 1096
QY	121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKSKTLPFLMQYSESE 180	QY 1081 KVEPPEVNPPLIIVGSSVGGLLALLALITAAALYKLGFPKROYKDMSESGSPGAEPPQ 1137
DB	137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKSKTLPFLMQYSESE 196	DB 1097 KVEPPEVNPPLIIVGSSVGGLLALLALITAAALYKLGFPKROYKDMSESGSPGAEPPQ 1153
QY	181 RIHFTFKFQNNPNPSLKIPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 240	
DB	197 RIHFTFKFQNNPNPSLKIPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 256	
QY	241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAPFSEKSRQELNTVASKPRDHVFOIN 300	
DB	257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVGDAPFSEKSRQELNTVASKPRDHVFOIN 316	
QY	301 NFALAKTIONQIREKIFALEGTGTGSSSPEHEMSOEGPSAITSNGPLLSVTGSDWAG 360	
DB	317 NFALAKTIONQIREKIFALEGTGTGSSSPEHEMSOEGPSAITSNGPLLSVTGSDWAG 376	
QY	361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAAAIILNRVQSLVLAGPYQHIGLVAMER 420	
DB	377 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAAAIILNRVQSLVLAGPYQHIGLVAMER 436	
QY	421 QNTGMWESNANVKTQIGYFAGSLCSVDVDSNGSTDLVLIGAPHYVYQTRGQSVCEPL 480	
DB	437 QNTGMWESNANVKTQIGYFAGSLCSVDVDSNGSTDLVLIGAPHYVYQTRGQSVCEPL 496	
QY	481 PRQQRARWQDAVLYGEQGPWRFGAALTIVLGDVNGDKLTDVAI GAPGEENRGAIVLFP 540	
DB	497 PRQQRARWQDAVLYGEQGPWRFGAALTIVLGDVNGDKLTDVAI GAPGEENRGAIVLFP 556	
QY	541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDLTVGAQGHVILLRSQ 600	
DB	557 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDLTVGAQGHVILLRSQ 616	
QY	601 PVLVRKAIENFNEPREVARNVFECNDQVWKGKAGEVRVCLVHVKSTRDLREGQIOSVVT 660	
DB	617 PVLVRKAIENFNEPREVARNVFECNDQVWKGKAGEVRVCLVHVKSTRDLREGQIOSVVT 676	
QY	661 YDLALDSGRPHSRAVNETKNSTRROTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNFP 720	
DB	677 YDLALDSGRPHSRAVNETKNSTRROTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNFP 736	
QY	721 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFENKQNDNICQDDLSITFSFMSLDCLVVG 780	
DB	737 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPPFENKQNDNICQDDLSITFSFMSLDCLVVG 796	
QY	781 GPREFNVTVTRNDGSDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWSLACESASTEV 840	
DB	797 GPREFNVTVTRNDGSDSVRTQVTFPPPLDLSYRKVSTLQNRQSRWSLACESASTEV 856	
QY	841 SGALKSTSCSINHPIIPKSEVTFNITFDVDSKASIGNKLLKANVTSENNNPRTNKTFP 900	
DB	857 SGALKSTSCSINHPIIPKSEVTFNITFDVDSKASIGNKLLKANVTSENNNPRTNKTFP 916	
QY	901 QLELPVKIAYVMVTVTSHGVSSTKYLNFNTAGENTSRRVMQHOVQVSNLQORSILPSLFLVFP 960	
DB	917 QLELPVKIAYVMVTVTSHGVSSTKYLNFNTAGENTSRRVMQHOVQVSNLQORSILPSLFLVFP 976	

RESULT 2
AAB07360
ID AAB07360 standard; protein; 1153 AA.
XX
AC AAB07360;
XX
DT 17-JAN-2001 (first entry)
XX
Human CD11b protein sequence.
DE
Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; leu-CM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; Crohn's disease;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11b.
XX
Homo sapiens.
OS
XX
WO200029446-A1.
PN
XX
25-MAY-2000.
PD
XX
16-NOV-1999; 99WO-US027139.
PF
XX
16-NOV-1998; 98US-00193043.
PR
XX
08-JUL-1999; 99US-00350259.
PT
XX
(ICOS-) ICOS CORP.
PA
XX
Gallatin NW, Van Der Vliet N;
PI
XX
WPI; 2000-387751/33.
DR
XX
Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
XX injury sites.
PS
XX
Example 5; Fig 1; 270pp; English.
CC
Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11b. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha d (AA60014 and AAB07359). The present sequence has
CC approximately 60% identity to the protein sequence of alpha d. The
CC Alpha d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease
XX

Query Match	99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity	99.3%; Pred. No. 0;
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;
QY	1 FNLDTENAMTFQENAGFGQSVVQLQSGVVVVGAPQEIIVAAQNRGSLVQCDYSTGSCBPI 60
DB	17 FNLDTENAMTFQENAGFGQSVVQLQSGVVVVGAPQEIIVAAQNRGSLVQCDYSTGSCBPI 76
QY	61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHOTCSSENTVYVKGCLTFLPSNLRQQPOK 120
DB	77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHOTCSSENTVYVKGCLTFLPSNLRQQPOK 136
QY	121 FPALRGCCPOEDSDIAFLVNDGSGIIPHDFRAKRPISITWEMQLKKSKTLPGLMOYSSEF 180
DB	137 FPALRGCCPOEDSDIAFLVNDGSGIIPHDFRKEKFEFVMTWEMQLKKSKTLPGLMOYSSEF 196
QY	181 RIHFTKEFQNNENPSRLIKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB	197 RIHFTKEFQNNENPSRLIKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY	241 TDGEKFGDPIGYEDVTPEADREGVIRVIVGVGDAPFSEKSRQELNATVASKPRDHFVQIN 300
DB	257 TDGEKFGDPIGYEDVTPEADREGVIRVIVGVGDAPFSEKSRQELNATVASKPRDHFVQIN 316
QY	301 NFEALKTIQNLREKIIFAEIGTQTGSSSFHEHMSQEGFSAATTSNGPLISTVGSYDWAG 360
DB	317 NFEALKTIQNLREKIIFAEIGTQTGSSSFHEHMSQEGFSAATTSNGPLISTVGSYDWAG 376
QY	361 GVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAIILNRVQSIVLGPARYQHIGLVAMFR 420
DB	377 GVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAIILNRVQSIVLGPARYQHIGLVAMFR 436
QY	421 QNTGWMESNANVKQTQIGAVFGSLCSVDVDSNGSTDVILGAPHYVEOTRGQSVSVCPL 480
DB	437 QNTGWMESNANVKQTQIGAVFGSLCSVDVDSNGSTDVILGAPHYVEOTRGQSVSVCPL 496
QY	481 PRGORARQCDVILGYEQGPWGRFGAALTIVLGPVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB	497 PRGORARQCDVILGYEQGPWGRFGAALTIVLGPVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY	541 HGTSGSGISPSHSORIIAGSKLSPLQVFGQSLSGQDLTWDGLVLDITVGAQGHVLLRSQ 600
DB	557 HGTSGSGISPSHSORIIAGSKLSPLQVFGQSLSGQDLTWDGLVLDITVGAQGHVLLRSQ 616
QY	601 PVLVRKALMEFNPREVARNVECDNDVVVKGEAGEVRVCLHVQKSTRDLREQQIQSVVT 660
DB	617 PVLVRKALMEFNPREVARNVECDNDVVVKGEAGEVRVCLHVQKSTRDLREQQIQSVVT 676
QY	661 YDLALDSGRPHSRAVFNNTKSTRQTQVILGLOTCTETLKLQIPNCIEDPVSIVLRLNF 720
DB	677 YDLALDSGRPHSRAVFNNTKSTRQTQVILGLOTCTETLKLQIPNCIEDPVSIVLRLNF 736
QY	721 SLVGTPLSAGNLRPVLADQRLETFALPPEKKCGNDNICQDDLSITTFPSMLDCLVWG 780
DB	737 SLVGTPLSAGNLRPVLADQRLETFALPPEKKCGNDNICQDDLSITTFPSMLDCLVWG 796
QY	781 GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRQORSWRILACSESASTEV 840
DB	797 GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRQORSWRILACSESASTEV 856
QY	841 SGALKSTSCSINHPIPPENSEVTFINITDQVDSKASIGNKLLKANVTSNNMPTNKTEF 900
DB	857 SGALKSTSCSINHPIPPENSEVTFINITDQVDSKASIGNKLLKANVTSNNMPTNKTEF 916
QY	901 QLELPVKYAVVWVITSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVPV 960
DB	917 QLELPVKYAVVWVITSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVPV 976
QY	961 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVWNCISIAQRIQCDIP 1020
DB	977 RLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVWNCISIAQRIQCDIP 1036

```
XX SQ Sequence 1153 AA;
Query Match 99.8%; Score 5862; DB 5; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARFGQSVVOLQGRVVRVVGAPQBIIVAAVNRGSLYQCDYSTGSCPEPI 60
Db 17 FNLDTENAMTFQENARFGQSVVOLQGRVVRVVGAPQBIIVAAVNRGSLYQCDYSTGSCPEPI 76

Qy 61 RLQVPEAVNNSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOPK 120
Db 77 RLQVPEAVNNSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOPK 136

Qy 121 FPEALRGCPQEDSDIAFLVDGSGIIPHDPRAREFTSTVMEQLKKSKTLFSLMQYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGIIPHDPRAREFTSTVMEQLKKSKTLFSLMQYSSEF 196

Qy 181 RIHPTFKEFONNPRSLKIPITQGLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
Db 197 RIHPTFKEFONNPRSLKIPITQGLGRTHATGIRKVVRELENTINGARKNAFKILWI 256

Qy 241 TDGEKFGDPLGYEDVTPADREGVIRVYVGVGDAFRSEKSRQBLNTVASKPPDRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVTPADREGVIRVYVGVGDAFRSEKSRQBLNTVASKPPDRDHVFQYN 316

Qy 301 NPEALKTIQNLREKIPAIETGTQSGSSSPEHEMSQEGFSAATISNGPILSTVGSVDWAG 360
Db 317 NPEALKTIQNLREKIPAIETGTQSGSSSPEHEMSQEGFSAATISNGPILSTVGSVDWAG 376

Qy 361 GVFLYTSKESKSTFTNTRVDSNDNDAYLVGAAAIILNRNVQSLVGLGAPRYQHILGVAMFR 420
Db 377 GVFLYTSKESKSTFTNTRVDSNDNDAYLVGAAAIILNRNVQSLVGLGAPRYQHILGVAMFR 436

Qy 421 QNTGMESNANVGTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYVETRGQGVQVCPBL 480
Db 437 QNTGMESNANVGTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYVETRGQGVQVCPBL 496

Qy 481 PRQORARWQCDAYLYEGOGQPMRFGAALTVDGVNAGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRQORARWQCDAYLYEGOGQPMRFGAALTVDGVNAGDKLTDVAIGAPGEEDNRGAVILF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTVGAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDTVGAQGHVLLLRQ 616

Qy 601 PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTDRBLREGIOQSVVT 660
Db 617 PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTDRBLREGIOQSVVT 676

Qy 661 YDLALDSGRPHSRAVFNETHKNSTRRTQVILGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETHKNSTRRTQVILGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736

Qy 721 SLNGTPLSAPGNLRPLVLAEDAQLFTALPPEKNCNDNICQDDLSITFSFMSLCLVVG 780
Db 737 SLNGTPLSAPGNLRPLVLAEDAQLFTALPPEKNCNDNICQDDLSITFSFMSLCLVVG 796

Qy 781 GPREFNVTVVRMDGSDSVRTQVTFPPFLDLSYRKVSTLQNGRSQSRSLACESASSTEV 840
Db 797 GPREFNVTVVRMDGSDSVRTQVTFPPFLDLSYRKVSTLQNGRSQSRSLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSVASLGNKLLKANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSVASLGNKLLKANTYSENMPRTNKTFF 916

Qy 901 QLELPVKYAVYVTVSHGVSTKYNTFASNTSRVHQYQVSNLQORSLPLSLVPLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYNTFASNTSRVHQYQVSNLQORSLPLSLVPLVPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1036
```

RESULT 4

ABG61469
ID ABG61469 standard; protein; 1153 AA.

XX AC ABG61469;

XX DT 27-AUG-2002 (first entry)

XX DE Human Beta2 integrin alphaCD11b subunit.

XX KW Beta2 integrin; alphaD subunit; CD11c subunit; LAD; leukocyte adhesion deficiency; inflammatory response; diabetes; multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective; inflammatory bowel disease; Crohn's disease; ulcerative colitis; immune complex alveolitis; leukaemia; ICAM-2; VCAM-1; anti-inflammatory; intracellular cell adhesion molecule; vascular cell adhesion molecule; locomotor recovery; locomotor damage; locomotor impairment; autonomic dysfunction; sensory dysfunction; spinal cord injury.

OS Homo sapiens.

PN WO200230980-A2.

XX 18-APR-2002.

XX PF 15-OCT-2001; 2001WO-US032059.

XX PR 13-OCT-2000; 2000US-00688307.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van Der Vieren M;

XX DR WPI; 2002-463260/49.

XX PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury.

XX PS Example 5; Page 191-194; 270pp; English.

XX CC The invention relates to promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury by administering an anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha d CDNA and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia). The present sequence is a Beta2 integrin alpha subunit sequence included for comparison with the Beta2 integrin alpha subunit sequences

SQ Sequence 1153 AA;
 Query Match 99.8%; Score 5862; DB 5; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY	1	ENLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAAQNRGSLVQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAAQNRGSLVQCDYSTGSCPEI	76
QY	61	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK	120
DB	77	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK	136
QY	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRAKFPISTWEOLEKSKTFLSLMOYSEEP	180
DB	137	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRAKFPISTWEOLEKSKTFLSLMOYSEEP	196
QY	181	RHIFTFKQFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI	240
DB	197	RHIFTFKQFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILIVI	256
QY	241	TGDEKFGDPLGHEDVIPADREGVIRYVIGVDAPRSEKSKQELNTVASPRDHVQIN	300
DB	257	TGDEKFGDPLGHEDVIPADREGVIRYVIGVDAPRSEKSKQELNTVASPRDHVQIN	316
QY	301	NFSALKTIQNLREKIPALEGTQSGSSPFHEMSQSGFSAITNSGPLLSTWGSYDNAG	360
DB	317	NFSALKTIQNLREKIPALEGTQSGSSPFHEMSQSGFSAITNSGPLLSTWGSYDNAG	376
QY	361	GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAIILNRVQSLVILGAPRYOHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAIILNRVQSLVILGAPRYOHIGLVAMFR	436
QY	421	QNTGWESNANKYGTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYEOTRGGOVSCPL	480
DB	437	QNTGWESNANKYGTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYEOTRGGOVSCPL	496
QY	481	PRQGRARWCDDAVLYGEOQCPAGREGAALTVLGDNVNDKLTDAVIGAPGEDNKGAYILF	540
DB	497	PRQGRARWCDDAVLYGEOQCPAGREGAALTVLGDNVNDKLTDAVIGAPGEDNKGAYILF	556
QY	541	HGTSGSGIPSHSRIASGLSPRIQYQSGSLGGQDLTMDGLVDLTVGAGHVLILRSQ	600
DB	557	HGTSGSGIPSHSRIASGLSPRIQYQSGSLGGQDLTMDGLVDLTVGAGHVLILRSQ	616
QY	601	PVLRVKAIEMFNEPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
DB	617	PVLRVKAIEMFNEPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSRVNETKNSRTOQVILGTQCTETKLQLPNCIEDPSPVILRLNF	720
DB	677	YDLALDSGRPHSRVNETKNSRTOQVILGTQCTETKLQLPNCIEDPSPVILRLNF	736
QY	721	SLVGTPLSAFNLRLPVLAEDAQRLEALFPPEKNGNDNICODDLSTFPMSLDCLVWG	780
DB	737	SLVGTPLSAFNLRLPVLAEDAQRLEALFPPEKNGNDNICODDLSTFPMSLDCLVWG	796
QY	781	GPREFNVTTVNDEGDSYRTQVITFFFLDLSYRKVSTLQNRQSQRWELACSSASSTEV	840
DB	797	GPREFNVTTVNDEGDSYRTQVITFFFLDLSYRKVSTLQNRQSQRWELACSSASSTEV	856
QY	841	SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNTEP	900
DB	857	SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNTEP	916
QY	901	QLELPKVAVMVMTSHGVSTKYLNFNTASNTSRVWQHOYQVNLGORSIPISLFLVPV	960
DB	917	QLELPKVAVMVMTSHGVSTKYLNFNTASNTSRVWQHOYQVNLGORSIPISLFLVPV	976
QY	961	RLNQTVIMDRPQVTSSENLSSTCHTKERLPKSHSDFLAELRKAPVWVNCISVACQRIQDIP	1020
DB	977	RLNQTVIMDRPQVTSSENLSSTCHTKERLPKSHSDFLAELRKAPVWVNCISVACQRIQDIP	1036

Query Match 99.8%; Score 5862; DB 5; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLLPGQGAFFVRSOTET 108
 DB 1037 PFGIOEEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLLPGQGAFFVRSOTET 1096
 QY 1081 KYPEPEVNPLPLIVGSSVGGGLLALITAAALYKLGFFPKQYKDMSEGPPGAEPO 1137
 DB 1097 KYPEPEVNPLPLIVGSSVGGGLLALITAAALYKLGFFPKQYKDMSEGPPGAEPO 1153

RESULT 5
 AA014428
 ID AAO14428 standard, protein; 1153 AA.
 XX AAO14428;
 AC AAO14428;
 DT 03-MAY-2002 (first entry)
 XX Integrin Mac-1 alpha subunit.
 DE Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
 KW open conformation; integrin related inflammatory disorder;
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;
 KW viral infection; cancer; gene therapy; vaccine;
 KW bioactive agent screening.
 XX Unidentified.
 XX WO200204521-A2.
 PN 17-JAN-2002.
 XX 09-JUL-2001; 2001WO-US021805.
 PF 07-JUL-2000; 2000US-0216600P.
 PR (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX (BLAO-) CENT BLOOD RES.
 PA Springer T;
 XX WPI; 2002-148167/19.
 DR New integrin I domain protein having alteration in at least 2
 XX noncontiguous regions and exits in an open conformation, useful for
 PT treating, preventing or suppressing inflammatory or immunological
 PT disorders.
 PT Example 1; Fig 1F; 90pp; English.
 PS The invention comprises structurally biased variant integrin inserted (I)
 XX domain proteins, wherein the alterations to the protein occur in at least
 CC two noncontiguous regions. Specifically the variant integrin I domain
 CC proteins are structurally biased to exist in the open conformation,
 CC thereby altering the binding ability of the protein. The invention also
 CC comprises nucleic acids encoding the variant integrin I domain proteins.
 CC The integrin I domain proteins and nucleic acids are useful for treating,
 CC preventing or suppressing integrin related inflammatory and immunological
 CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
 CC proteins and nucleic acids can also be used for treating: ischaemia/
 CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
 CC infection; and cancer. The variant integrin I domain nucleic acids and
 CC proteins may be used in gene therapy, as vaccines and to screen for
 CC bioactive agents. The present amino acid sequence represents the Mac-1
 CC alpha subunit of integrin
 XX Sequence 1153 AA;
 SQ

QY 1 FNLDTENAMTFOENARFGQSVVOLQSGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVVOLQSGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVAVNMSLGLSLAATTSPOILLACGPTVHOTCSENTRYKGLCPFGSLRQPOOK 120
DB 77 RLOVPVAVNMSLGLSLAATTSPOILLACGPTVHOTCSENTRYKGLCPFGSLRQPOOK 136
QY 121 FPBALRGCPQEDSDIAFLVDGSGSIIPHPRRAKBFISTVMEQLKSKSLFSLMQYSEEP 180
DB 137 FPBALRGCPQEDSDIAFLVDGSGSIIPHPRRAKBFISTVMEQLKSKSLFSLMQYSEEP 196
QY 181 RIHFTPKBFQNNPNPSLKIPIQLLGRTHATGIRKVVRELNIINGARKNAFKILLI 240
DB 197 RIHFTPKBFQNNPNPSLKIPIQLLGRTHATGIRKVVRELNIINGARKNAFKILLI 256
QY 241 TDCEKFGDPLGYEDVPEADREGVIRVIGVGDFAFRSEKSRQELNIVASKPPDRHVFQIN 300
DB 257 TDCEKFGDPLGYEDVPEADREGVIRVIGVGDFAFRSEKSRQELNIVASKPPDRHVFQIN 316
QY 301 NFALKTIONOLREKIPAIISGTOTGSSSSPEHEMSQEGPSAALTNGPILLSVGSYDWAG 360
DB 317 NFALKTIONOLREKIPAIISGTOTGSSSSPEHEMSQEGPSAALTNGPILLSVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAIILNRVOSLVIGAPRYOHLGLVAMER 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAIILNRVOSLVIGAPRYOHLGLVAMER 436
QY 421 QNTGMWESNANVKGTGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCPIL 480
DB 437 QNTGMWESNANVKGTGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCPIL 496
QY 481 PRGRARWQCDVLYGQGGPWGFRGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVILF 540
DB 497 PRGRARWQCDVLYGQGGPWGFRGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVILF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVLDLVFGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLSGQDLTMDGLVLDLVFGAQGHVLLRSQ 616
QY 601 PVLRYKAIMFNPREVARNVPCNDQVVKGEAGEVAVCLHVOKSTEDRLREGOISQVVT 660
DB 617 PVLRYKAIMFNPREVARNVPCNDQVVKGEAGEVAVCLHVOKSTEDRLREGOISQVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSIRROQVLGLTQTCETLKLQPNCTEDPVSIVLRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSIRROQVLGLTQTCETLKLQPNCTEDPVSIVLRLNF 736
QY 721 SLVGTPLSAPGNLSPVLAEDQRLFTALPPEKCKGNDNICQDDLSITPFSMLDCLVVG 780
DB 737 SLVGTPLSAPGNLSPVLAEDQRLFTALPPEKCKGNDNICQDDLSITPFSMLDCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNTEP 900
DB 857 SGALKSTSCSINHPIFPENSSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNTEP 916
QY 901 QLELPVKYAVVMVTSYKLYNFTASENTSRVMOHYQVSNLQORSIPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSYKLYNFTASENTSRVMOHYQVSNLQORSIPISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVNVCSIAVCORIQDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVNVCSIAVCORIQDIP 1036
QY 1021 PFGIOEEFNATLKGNSLFDWIKTSHNHLILVSTABILFNDSVFTLLPGQCAFVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNSLFDWIKTSHNHLILVSTABILFNDSVFTLLPGQCAFVRSOTET 1096
QY 1081 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGSPGABPQ 1137

DB 1097 KVEPPEVFNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGSPGABPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX ADD25615;
DT 15-JAN-2004 (first entry)
XX Binding domain-immunoglobulin fusion protein-associated protein #85.
DE Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antirheumatic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX Unidentified.
OS US2003118592-A1.
XX 26-JUN-2003.
XX 25-JUL-2002; 2002US-00207655.
XX 17-JAN-2001; 2001US-0367358P.
XX 17-JAN-2002; 2002US-00053530.
XX 03-JUN-2002; 2002US-0385691P.
XX (GENE-) GENE-CRAFT INC.
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
PI WPI; 2003-801317/75.
XX New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX Disclosure; SEQ ID NO 176; 157pp; English.
XX The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin hinge
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
CC hinge region polypeptide, derived from (a) having 3 or more cysteine residues;
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
CC contains 2 cysteine residues, where the first cysteine is not mutated; a
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
CC (a) having 3 or more cysteine residues, where the mutated human IgG1
CC immunoglobulin hinge region polypeptide contains no more than one
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
CC polypeptide, derived from (a) having 3 or more cysteine residues; where
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
CC capable of at least one immunological activity comprising antibody
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
CC binding domain polypeptide is capable of specifically binding to an
CC antigen. Also included are an isolated polynucleotide encoding the
CC binding domain-immunoglobulin fusion protein, a recombinant expression
CC construct comprising the polynucleotide (operably linked to a promoter),
CC a host cell transformed or transfected with a recombinant expression
CC construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin
CC fusion protein or polynucleotide and a carrier, and treating a subject
CC having or suspected of having a malignant condition or a B-cell disorder.
CC The binding domain-immunoglobulin fusion protein is useful for treating a
CC subject having or suspected of having a malignant condition or a B-cell
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
CC sclerosis or autoimmune disease. The present sequence is a binding domain
CC -immunoglobulin fusion protein-associated protein sequence. Note: The
CC sequence data for this patent formed part of the printed specification
CC and is also available in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
CC identified the sequences in the printed specification by their SEQ ID
CC number therefore none of the sequences can be explicitly identified.

XX
SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 7; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPOEIVAAHQRSLSYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPOEIVAAHQRSLSYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLTFLGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLTFLGSNLRQPOK 136

QY 121 FPEARLGCPOEDSLAFLVDGSGSIIPHDFRAKEFISTVWOLKSKTLFLSLMOYSEF 180
DB 137 FPEARLGCPOEDSLAFLVDGSGSIIPHDFRANKFVSTWELQKSKTLFLSLMOYSEF 196

QY 181 RIHFTPFKEFQNNPNRSLKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
DB 197 RIHFTPFKEFQNNPNRSLKPIITOLLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256

QY 241 TIGEXFGDPLGYEDVITPEADRGVIRYVIGVGDAFERSKSRQELMTVASKPRDHFVFOIN 300
DB 257 TIGEXFGDPLGYEDVITPEADRGVIRYVIGVGDAFERSKSRQELMTVASKPRDHFVFOIN 316

QY 301 NFEALKTIQNLREKIPALFEGTQTSSESSFEHMSQEGFSAITSNGLPSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIPALFEGTQTSSESSFEHMSQEGFSAITSNGLPSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTPINMTRVDSMDNAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMPR 420
DB 377 GVFLYTSKSKSTPINMTRVDSMDNAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 496

QY 481 PRQQRARWQCDVLYGEOQPMGRFGAALTIVIGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQQRARWQCDVLYGEOQPMGRFGAALTIVIGVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLSGGQDLTMDGLVLTGCAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLSGGQDLTMDGLVLTGCAQGHVILLRSQ 616

QY 601 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHYQKSTRDLRLEGQIQSVVT 660
DB 617 PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHYQKSTRDLRLEGQIQSVVT 676

QY 661 YDLALDSGPHSRVAVNETKSTRCTQVLGTOTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGPHSRVAVNETKSTRCTQVLGTOTCETLKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPFFKNCNDNICQDDLSITFFSMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPPFFKNCNDNICQDDLSITFFSMSLDCLVVG 796

QY 781 GPREFNVTVTRMDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQSRWRLACBSASSTEV 840
DB 797 GPREFNVTVTRMDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQSRWRLACBSASSTEV 856

QY 841 SGALKSTSCSINHPFIPENSEVTNITFDVDSKASLGNKLLIKANVTSENKMPSTNKTET 900
DB 857 SGALKSTSCSINHPFIPENSEVTNITFDVDSKASLGNKLLIKANVTSENKMPSTNKTET 916

QY 901 QLELPVXYAVVWVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVXYAVVWVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSLPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFPSNLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFPSNLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQIQCDIP 1036

QY 1021 PFGIOEESFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSOTET 1080
DB 1037 PFGIOEESFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSOTET 1096

QY 1081 KVSEPFVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKQDMSEGGPPGABPQ 1137
DB 1097 KVSEPFVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKQDMSEGGPPGABPQ 1153

RESULT 7
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
XX AC AAR04136;
XX DT 25-MAR-2003 (revised)
XX DT 07-SEP-1990 (first entry)
XX DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX KW non-specific defence system; integrin gene superfamily.
XX OS Synthetic.
XX PH Key
XX Region 1..16
XX Location/Qualifiers
XX /label= signal peptide
XX Modified-site 86..88
XX /label= putative N-glycosylation site
XX Modified-site 240..242
XX /label= putative N-glycosylation site
XX Modified-site 391..393
XX /label= putative N-glycosylation site
XX Modified-site 469..471
XX /label= putative N-glycosylation site
XX Modified-site 693..695
XX /label= putative N-glycosylation site
XX Modified-site 697..699
XX /label= putative N-glycosylation site
XX Modified-site 735..737
XX /label= putative N-glycosylation site
XX Modified-site 802..804
XX /label= putative N-glycosylation site
XX Modified-site 881..883
XX /label= putative N-glycosylation site
XX Modified-site 901..903
XX /label= putative N-glycosylation site
XX Modified-site 912..914
XX /label= putative N-glycosylation site
XX Modified-site 941..943
XX /label= putative N-glycosylation site
XX Modified-site 947..949
XX /label= putative N-glycosylation site
XX Modified-site 979..981
XX /label= putative N-glycosylation site
XX Modified-site 994..996

```

FT Modified-site /label= putative N-glycosylation site
FT 1022..1024
FT /label= putative N-glycosylation site
FT 1045..1047
FT Modified-site /label= putative N-glycosylation site
FT 1051..1053
FT Modified-site /label= putative N-glycosylation site
FT 1076..1078
FT Modified-site /label= putative N-glycosylation site
FT 1106..1134
FT Region /label= putative transmembrane region
XX EP364690-A.
XX 25-APR-1990.
XX 17-AUG-1989; 89EP-00115159.
XX 23-AUG-1988; 89US-00235353.
XX 09-MAR-1989; 89US-00321239.
XX (DAND ) DANA FARBER CANCER INST INC.
XX Springer TA, Corbi A;
XX WPI: 1990-125938/17.
XX N-PSDB; AAQ04043.
XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
XX inflammation and viral infections, and in diagnosis.
XX Disclosure; Page 2; -pp; English.
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
XX recognition of and migration to sites of inflammation. It also attaches
XX to cellular substrates as part of this function making it useful in
XX visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
XX superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
XX -MAR-2003 to correct PA field.)
XX SQ Sequence 1153 AA;

Query Match 99.6%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLOVPVEANVMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
DB 77 RLOVPVEANVMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEOLKESKTLFSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEOLKESKTLFSLMOYSEEP 196

QY 181 RIETFKFQNNPRSLKPIITQLGRTHATGIRKVVRELFTNITNGARKNAFKILILI 240
DB 197 RIETFKFQNNPRSLKPIITQLGRTHATGIRKVVRELFTNITNGARKNAFKILIVI 256

QY 241 TDGEKGDPLGYEDVTPRADREGVIRVIGVGAPRSEKSRQELNTVASKPRDHDVQFN 300
DB 257 TDGEKGDPLGYEDVTPRADREGVIRVIGVGAPRSEKSRQELNTVASKPRDHDVQFN 316

QY 301 NFALKTIQNLREKIFAIBGTQTSSSSPHEMSQEGFSAATISNGPLLTSTVGSYDNAG 360
DB 317 NFALKTIQNLREKIFAIBGTQTSSSSPHEMSQEGFSAATISNGPLLTSTVGSYDNAG 376

QY 361 GVFLYTSKEKSTINTRVDSNDNDAYLVGAAAIILNRVQSLVLCAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTINTRVDSNDNDAYLVGAAAIILNRVQSLVLCAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYTEQTRGGQSVUCLP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYTEQTRGGQSVUCLP 496
QY 481 PRGORARWQCDVLYGQGGPWGRFGAALTIVLGDVNDKDLTDVAIGAPGEEDNRGAVYLP 540
DB 497 PRGORARWQCDVLYGQGGPWGRFGAALTIVLGDVNDKDLTDVAIGAPGEEDNRGAVYLP 556
QY 541 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLGQGLTMDGLVDLTVGAGHVLRLRSQ 600
DB 557 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLGQGLTMDGLVDLTVGAGHVLRLRSQ 616
QY 601 PVLRVKALMBENPREVARNVPECDNDQVVKGEACEVCLFHVQKSTRDRLEGOIQSVVT 660
DB 617 PVLRVKALMBENPREVARNVPECDNDQVVKGEACEVCLFHVQKSTRDRLEGOIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAPGNLFPVLAEDAORLFTALFPPEKNCNDNICODDLSTIFSFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLFPVLAEDAORLFTALFPPEKNCNDNICODDLSTIFSFMSLDCLVVG 796
QY 781 GPREFNVTIVRNDEGDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWSRLACESASSTEV 840
DB 797 GPREFNVTIVRNDEGDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWSRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGKLLKLLKANTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGKLLKLLKANTSENMMPTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCTKRLPSHSDPLAELKAPVYVNCISIAVCRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCTKRLPSHSDPLAELKAPVYVNCISIAVCRIQCDIP 1036
QY 1021 PFGIOEBENATLKNLSDFDWYIKTSHNHLIVSTAELFNDVSFTLLPQCGAFVRSQTET 1080
DB 1037 PFGIOEBENATLKNLSDFDWYIKTSHNHLIVSTAELFNDVSFTLLPQCGAFVRSQTET 1096
QY 1081 KVEPPEVNPPLPIVGVSGGLLILALITAAALYKLGPPKROYKDMMSGGPPGABPQ 1137
DB 1097 KVEPPEVNPPLPIVGVSGGLLILALITAAALYKLGPPKROYKDMMSGGPPGABPQ 1153

RESULT 8
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1991 (first entry)
XX
DE p150.95 alpha subunit encoded by clone lambdaX47.
XX
KW p150,95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
XX rhinovirus.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..19 /label= signal peptide
FT FT Region 20..44 /label= N-terminus
FT FT Modified-site 61..63 /label= glycosylation site

```


KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
OS Homo sapiens.
XX US5728533-A.
XX 17-MAR-1998.
XX 07-JUN-1995; 95US-00485618.
XX 23-DEC-1993; 93US-00173497.
XX 05-AUG-1994; 94US-00286889.
XX 21-DEC-1994; 94US-00362652.
XX (ICOS-) ICOS CORP.
XX Van Der Vieren M, Gallatin WM;
XX MPI; 1998-206565/18.
XX Screening assay for modulators of integrin binding - using immobilised or
XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.
XX Example 5; Fig 1A-D; 106pp; English.
XX This sequence represents a human beta-integrin CD11c subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
XX asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
XX and rheumatoid arthritis
XX Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity 60.8%; Pred. No. 8e-279;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDENAMTQENARCGQSVVQLQGSRRVVGAPQEIIVANQSGSYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAERVDASGFGDSVVGQYANSWVVGAPQEIIVANQSGSYQCDYSTGSCPEI 79
QY 61 RLQVPEAVNMVSLGLSLAATTPPQLACGPTVHTQCSNTYKGLCLFLGNSLRQPPQK 120
DB 80 GLQVPEAVNMVSLGLSLAATTPPQLACGPTVHTQCSNTYKGLCLFLGNSLRQPPQK 137
QY 121 PPEALRGCPQEDSIAFLVDSGSIIPHDPERAKEFIPTWQELKSKTLFSLMOYSEEP 180
DB 138 LPVRQECFQEQDIFVLDGSGSSISNFMATMNFVRAVISQFQRETFSLMQFNKP 197
QY 181 RIHFTFEFQNNPRLSLIKPITQILGRTHTATGIRKVRFLNITNGARKNAPKILILI 240
DB 198 QHTFTFEFRTSNELSLASVHQLQGYTATATQNVVHRLFEASTGARRDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGURVYVGDFAFRKSRQRLNTYASKPPRDHVFQIN 300
DB 258 TDGKKGESLDYKDVIPWADAAGIRVYVGLFAPNRNNSWKELNDIAKSPQOEHPKVE 317
QY 301 NFEALKTIQNLREKIPAEIGTQGTGSSSPHEMSQEGFSAATISNGPLLSVGSYDWAG 360
DB 318 DFDALKDILQNLKEKIPAEIGTETISSSSFELEMAQEGFSAVFTPDGVPVGLGAVCSFTMSG 377
QY 361 GVFLYTSKEKSTFFINMTVDSMDNDVILGYAAAIIILNRVQSLVILGAPRYQHIGLVAMFR 420

DB 378 GAFLYPPNNSPTFFINMSQENVMDRDSYLGYSSTELALMKGVQSLVLAGAPRYQHIGKAVIFI 437
QY 421 QNTGHWENANVKGTOICAYFGASLCSVDVDSNGSTDLVILGAPHYETGTGGQVSVCP 480
DB 438 QVSRQMRKAEVIGTQIGSYFGASLCSVDVDSNGSTDLVILGAPHYETGTGGQVSVCP 497
QY 481 PRGQARWQCDVLYGEOGPWGRGAALTVLGDVNGDKLTDVAIGAPGEEDNRCGAVL 540
DB 498 PRGWR-RWNCDAVLYGEOGPWGRGAALTVLGDVNGDKLTDVWIGAPGEEDNRCGAVL 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLDLTVGAQHVLILRSQ 600
DB 557 HGVLGFSISPSHSQRIAGSKLSSRLQYFGQALSGQDLTMDGLVLDLTVGAQHVLILRSQ 616
QY 601 PVLRYKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVOKSTDRLEREGQIQSVVT 660
DB 617 PVLWGVSKQFIPABIPRSAFEQVYVSEQTLVQSNICLYIDKSKNLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVFNFTKSTRTQOTVLGTQTCETLKLQLPNCIEDPVSIVLRNLP 720
DB 677 LOLALAPGRLSPRAFQETKNSLSRVRVLGKAECENFNLLPSCVEDSVIPIILRNLP 736
QY 721 SLVGTPLSAGNLRPVLAEDAQRLFTALPPPEKNGNDNICODLSITFSFMSLDCLVVG 780
DB 737 TLVGPXLLAFRLRPLMLAALAOYFTASLPFFKNGADHICQDNLGISFSPFLKSLVVG 796
QY 781 GPREFNVITVENDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRMSWPLACESASSTEV 840
DB 797 SNLELNABVWVWNGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRMSWPLACESASSTEV 854
QY 841 SGALKSTSCINHPFPENSEVTFNITVDVSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 855 SQTWSTSCINHLIFRGAQITFLATFDVSPRAVLDRLLLIANVSSNNIPRTSKTIP 914
QY 901 QLELPKYAVYVWVTSHGVSSTYKLNFTAS-ENTSRVMOHQYQVSNLQGRSLPISLVLP 959
DB 915 QLELPKYAVYVWVTSHGVSSTYKLNFTAS-ENTSRVMOHQYQVSNLQGRSLPISLVLP 974
QY 960 VRLNQTVIWDROVTFSENSLSTCHKEPLPHSDFAELRKAQVYVNCIAYCQICQDI 1019
DB 975 VELNQAVWMDVEVSHQPNSLCSSEKTAIPASDPLAHIQNPNVLDSCIAGLRPRCDV 1034
QY 1020 PFGIOEBFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVSQTE 1079
DB 1035 PFSVQVEELDFTLKGNLSEGWVRQILQKXSVSVVAEIIIFDTSVYSQLPQGAFAVRAQTI 1094
QY 1080 TKVPEPVPNPPLIYVSSVGGILLALITALYALYKLGFFKQYKMMSE 1128
DB 1095 TVLEKYKHNPILYVSSIGILLALITALYALYKLGFFKQYKMMSE 1143

RESULT 10
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
XX W0200029446-A1.
XX

25-MAY-2000.
 16-NOV-1999; 99NO-US027139.
 16-NOV-1998; 98US-00193043.
 08-JUL-1999; 99US-00350259.
 (ICOS-) ICOS CORP.
 Gallatin Mx, Van Der Vieren M;
 WPI; 2000-387751/33.
 Use of novel anti-alpha integrin d monoclonal antibodies to inhibit macrophage infiltration and reduce inflammation at central nervous system injury sites.
 Example 5; Fig 1; 270pp; English.
 Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AA60014 and AA607359). The present sequence has approximately 66% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease
 Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 3; Length 1163;
 Best Local Similarity 60.8%; Pred. No. 8e-279;
 Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGRVVVGAQOEIVAAQORGLYQCDYSTGSCBPI 60
 DB 20 FNLDTEELTAFRVDSAGFSDVVQVANSVVVGAQOEIVAAQORGLYQCDYSTGSCBPI 79
 QY 61 RLQVPEAVNMSLGLSLAATSPDQLLACGPPVHQTCSNTYVKGCLFPLFQSNLRQOPQK 120
 DB 80 GLQVPEAVNMSLGLSLAATSPDQLLACGPPVHQTCSNTYVKGCLFPLFQSNLRQOPQK 137
 QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTWBEQAKSKTLFSLMOYSEF 180
 DB 138 LPVSRQECRQCDIVFLIDGSGSISSRNFAFMVFAVRAVISQFQPTQSLAQFSNKF 197
 QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTNGARKNAFKILILI 240
 DB 198 QHFTFEEFRSTNPLSLASVHLQGLQFTYTATQNVVHRLFAHSGARDAIKILIVI 257
 QY 241 TQEKFGDPLGYEDVTPEDRREGVIRYVGVCDAPRSEKSRQELNTVASKPRDHVQJIN 300
 DB 258 TDGKXEGSDLYKDVFPADADAGIIRYATGVGLAFQNRNSWKELNDIAKSPQEHIFVE 317
 QY 301 NFEALNTIQNLREKIPALEGTQTSSSSFEHMSQEGFSAAITSNGFLLSTVGSYDNAG 360
 DB 318 DFDALKDQNLKEKIPALEGTQTSSSSFEHMSQEGFSAAITSNGFLLSTVGSYDNAG 377
 QY 361 GVFLYTSKEKSTFNNTRVDSNDNAYLCYAAAIILRNVRQSLVLTGAPRYQHIGLVAMFR 420
 DB 378 GAFLYPPNNSPTFNNNSQENVDNRDSYLTALMKGVQSLVLTGAPRYQHIGLVAMFR 437
 QY 421 QNTGMWESNANVKGTOIGVAFGASLCSVDVDSNGSTDLVLICAPHYYEQTRGGQVSVCP 480

Db QY 438 QVSRQWRKAEVIGTQTIGSYFGASLCSVDVDTGSLDVLIGAPHYYEQTRGGQVSVCP 497
 QY 481 PRGORARWODAVLYGBQOPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
 Db 498 PRGWR-RWCDVLYGEGQHPWGRFGAALTVLGVDVNGDKLTDVWIGAPGEENRGAVYLF 556
 QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFQGSLSGGDLTMDGLVDLTGAGGHVLLRSQ 600
 Db 557 HGVLPSPISPSHSQRTAGSKLSPRLQYFQGSLSGGDLTMDGLVDLTGAGGHVLLRSQ 616
 QY 601 PVLRYKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTDRLERGQIQSVVT 660
 Db 617 PVLWVGSMQFIPABIPRPAFECRQVQVSEQTIVQSNICLYIDKRSKNLGSRLQSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETNSTRRTQVILGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
 Db 677 LDIALAPGRISPRAIPOETKNSLSRVRLGLKACENFNLLPSCVEDSVIPIILRLNF 736
 QY 721 SLVCTPLSAFENLPVLAEDAQRFTALFPKKNCGNDNICODDLSTESFMSLDCLVVG 780
 Db 737 TLVKKPLLAFLNCRPMLAALQRYFTASLPFKKNCADHICQDNILGISFSPGKLSLLVG 796
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 Db 797 SNLENAEVMVWMDGDSYGTITTFSPAGLSYRYVAEGQKQQLRSLHLC--CSAPVG 854
 QY 841 SGALKSTCSINHPFPENSEVFNITFDVDSKASLGNKLLKANTVSENNMPTNKTEP 900
 Db 855 SQGTWSTSCINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENINPTSKTIF 914
 QY 901 QLELPVYAVVYVTVSHGVSTKILNPTAS-ENTSRVMOHOVQVSNLQORSIPISLVLPV 959
 Db 915 QLELPVYAVVYVTVSHGVSTKILNPTAS-ENTSRVMOHOVQVSNLQORSIPISLVLPV 974
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAEIRKAPVNCSTAVQRIQCDI 1019
 Db 975 VELNQBAVMVMDVEVHPQNPQLRCSSEKIAPPASDPLAHQKQNPVLDSCSIAGCLRFCDV 1034
 QY 1020 PFGIOEENATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQAFVRSQTE 1079
 Db 1035 PPSVQBELDFTLKNLSFGWVRLQKCVSVSVVAELIIFDTSVYSQLPQGEAFRAQTI 1094
 QY 1080 TKVFFFPVNPPLIVGSSVGGLLLIITAAALYKLGFFKRYQKDMSE 1128
 Db 1095 TVLEKYVKNPIPLIVGSSIGLLLIITAAALYKLGFFKRYQKDMSE 1143

RESULT 11
 ABG61470
 ID ABG61470 standard; protein; 1163 AA.
 XX AC ABG61470;
 XX DT 27-AUG-2002 (first entry)
 XX DE Human Beta2 integrin alphaCD11c subunit.
 XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICM-R; VCM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX WO200230980-A2.
 XX PD 18-APR-2002.
 XX


```

PF 15-OCT-2001; 2001WO-US032059.
XX
PR 13-OCT-2000; 2000US-00688307.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Van Der Vieren M;
XX
XX WPI; 2002-463260/49.
XX
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
XX recovery, inhibiting locomotor damage, limiting locomotor impairment, or
XX limiting autonomic and sensory dysfunction following spinal cord injury.
XX
XX Example 5; Page 194-198; 270pp; English.
XX
XX The invention relates to promoting locomotor recovery, inhibiting
XX locomotor damage, limiting locomotor impairment, or limiting autonomic
XX and sensory dysfunction following spinal cord injury by administering an
XX anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
XX spinal cord injury victim. The method also involves the use of a ligand
XX selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,
XX vascular cell adhesion molecule). The method is useful for promoting
XX locomotor recovery, inhibiting locomotor damage, limiting locomotor
XX impairment, or limiting autonomic and sensory dysfunction following
XX spinal cord injury. In particular, the spinal cord injury comprises
XX compression of the spinal cord. The antibodies are also useful for
XX reducing inflammation at the site of a central nervous system injury. The
XX specification also details the identification of Beta2 integrin alpha d
XX cDNAs and proteins, for use in raising the antibodies. Beta2 integrins
XX are implicated in diseases such as IAD (leukocyte adhesion deficiency,
XX inflammatory response, diabetes, multiple sclerosis, arthritis, graft
XX atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative
XX colitis, immune complex alveolitis and leukaemia. The present sequence is
XX a Beta2 integrin alpha subunit sequence included for comparison with the
XX Beta2 integrin alpha d protein sequences
XX
XX Sequence 1163 AA;
XX
XX Query Match 58.6%; Score 3446; DB 5; Length 1163;
XX Best Local Similarity 60.8%; Pred. No. 8e-279;
XX Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
XX
XX 1 FNLDTEVAMTFQENARFGQSVVLQGSRRVVGAPQEIIVAAVQNGSLYQCDYTGSCPEI 60
XX 20 FNLDTELTAFRVDASAGFSDVQVYANSWVVGAPQKILAAVQIGGLYQGYTGACEPI 79
XX
XX 61 RLQVFEAVNMSGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
XX 80 GLQVPPAVNMSGLSLASTTSPQLLACGPTVHHEGGRNMYLTGLCFLGPT--QLTQR 137
XX
XX 121 FPSALRCQPEDSDIAFLVDCGSGSIIIPHDPRRAKEFISTVNEQLKSKSLFSLMOYSEEP 180
XX 138 LPVSRQRCPRQEQDIVFLIDGSGSISRRNFATWNVFRAVISQFRPSTQFSLMQFSNKF 197
XX
XX 181 RIHETPEFQNNPNRSLIKPITOLLGRTHATGIRKVRLEFNITNGASKNAPKILILI 240
XX 198 QTHETFEFRTSNPLSLASVHQLQOFTTATAIGNVRLPHASVGAARDAILKILIVI 257
XX
XX 241 TDGKFGDPLGYEDVIEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVPQIN 300
XX 258 TDGKKGDSLDYKDVPMDAAGIIRYVIGVLAQWRNSWKEINDIASRPSQEHIFKVE 317
XX
XX 301 NFEALKTQICLRKEIFALIGTQSGSSSEHEHMSOEGFSAATTSNCPLLSTVSGSDWAG 360
XX 318 DFDALKDIOQLKEKIFALEGTETETISSSSELEMAOQEGFSAVTFDGPVLGAVGFSWGS 377
XX
XX 361 GVFLYTSKESKSTFTNMTRVDSMDNDAYLGAAYAILNRRVQSLVGLAPRYOHIGLVAMER 420
XX 378 GAFLYPPNVSPTFNMSQENVMDSDYLGSTELALWKGVSILVGLAPRYOHIGKAVIFI 437
XX
XX 421 QNTGWESNANVKGTQICATFGASLCSVDVDSNGSDTLVIGAPHYYEQTGRGQSVSCPL 480
XX

```

```

Db 438 QVSRQWRKAEVIGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEQTGRGQSVSCPL 497
Qy 481 PRGQARWOCDAVLYGEGQPMGRFGAALTVLGDVNGDKLTDLVAIGAPGEENRGAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEGQPMGRFGAALTVLGDVNGDKLTDLVAIGAPGEENRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLTVCAGQGHVLLLRQ 600
Db 557 HGVLPSPISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLTVCAGQGHVLLLRQ 616
Qy 601 PVLRYKAIMEFNPREVARNVFCNDQVYKGEAGEVRVCLHVQKSTRDLRLESGQIQSVVT 660
Db 617 PVLWVGVSQMFIPASIPRSAPFCRBQVSEQLVQSNICLYIDKSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTVGLGTQCTETLKLQLPNCIEDPVPVILRLNF 720
Db 677 LDLALAPGKLSPLAFQETKNSLSKRVRLGKAHCENFNLLPSCVEDSVPIILRLNF 736
Qy 721 SLVGTPLSAPGNLRFVLAEDQRLFTALFPPEKNGCNDNICODDLSITFSFMSLCLVVG 780
Db 737 TLVGKPLLAFLNLRPMLAALAQRYFTASLPPEKNGCADHICQDNLGIGSFPGLKSLVG 796
Qy 781 GPREFNVTYVNDGEDSVRTQVTFPFLDLSYKRVSTLQNSORSORSWELACESASSTEV 840
Db 797 SNLELNAEYVWVNDGEDSYGTITITFSHPAGUSYRYVABGQKQGLRSLHLC--CSAPVG 854
Qy 841 SGALAKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXAVTSENMMPTNKTTF 900
Db 855 SGTWSTSCRINHLIFRGGAQITFLATFDSKAVGLDRLILLIANVSSNNIPTRTKTF 914
Qy 901 QLELPKVAIVMVTSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQBSIFLSLVFLVP 959
Db 915 QLELPKVAIVTVSGSHQFTKYLNFSESEKESHVAMERYQVNNLQGRDLQSVSNFWVP 974
Qy 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPQSHSDFLAELRKAPVNVNCSIAVCQITQCDI 1019
Db 975 VELNQEAVMVDEVSHPQNPSPSLRCSSEKIAPASDFLAHQKPNVLDCSIACCLAFRCDV 1034
Qy 1020 PFGQOEPNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDQSVFTLLPQCGAFVRSQTE 1079
Db 1035 PSFSVQEBLDFTKENLSPGWVRQILQKKVSVVSAEIIFTDTSVZSOLPQCGAFVRAQTI 1094
Qy 1080 TKVEPEVNPPLVGVSGVGLLLALITAAALYKLGFFKRYKQMMSE 1128
Db 1095 TVIERKYHNPPLVGVSGVGLLLALITAAALYKLGFFKRYKEMME 1143
XX
XX RESULT 12
XX ABU07406
XX ID ABU07406 standard; protein; 1163 AA.
XX AC ABU07406;
XX DT 28-JAN-2003 (first entry)
XX XX
XX DE Protein differentially regulated in prostate cancer #9.
XX XX
XX KW Prostate cancer; gene expression; differential regulation;
XX cancer marker; drug target; cancer detection; cancer diagnosis;
XX cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX OS Homo sapiens.
XX PN WO200281638-A2.
XX XX
XX PD 17-OCT-2002.
XX XX
XX PF 08-APR-2002; 2002WO-US010824.
XX XX
XX PR 06-APR-2001; 2001US-0281731P.
XX PR 06-APR-2001; 2001US-0281732P.
XX XX
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

```

XX Sun Z, Jay G;
XX WPI; 2003-058520/65.
XX Novel genes which are differentially regulated in prostate cancer, useful
XX for diagnosing prostate cancer in prostate tissue sample and assessing
XX therapeutic or preventive intervention in prostate cancer patients.
XX Claim 1; Page 225-228; 416pp; English.
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially regulated in prostate cancer.
XX Preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially regulated in prostate cancer
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX the test agent to modulate a biological activity of the polypeptide, and
XX determining whether the test agent modulates the biological activity. (I)
XX is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX preventing or treating, determining predisposition to diseases and
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used for expressing the polypeptide and thus
XX for searching specific binding partners of the polypeptide. (I) is useful
XX in therapeutic applications to treat prostate cancer. The identification
XX of specific genes, and groups of genes, expressed in pathways
XX physiologically relevant to prostate cancer permits the definition of
XX functional and disease pathways and the delineation of targets in these
XX pathways which are useful in diagnostic, therapeutic, and clinical
XX applications. This is the amino acid sequence of a protein differentially
XX regulated in prostate cancer
XX
XX Sequence 1163 AA;
Query Match 58.6%; Score 3444; DB 6; Length 1163;
Best Local Similarity 60.7%; Pred. No. 1.2e-278;
Matches 585; Conservative 139; Mismatches 299; Indels 6; Gaps 4;
Qy 1 FNLDTENAFQENARGFGQSVVLOGSRVWVVGAPQEIYAANORGSLYOCXYSTGSCPT 60
Db 20 FNLDTEELTAFRVDSAGFGVGVYVYANVWVVGAPQKITAAQNTGGYOGYSTGACEPI 79
Qy 61 RLQVPEAVNMNLSGLSIAATTSPPQLLAGQPTVHQTCSNTYVKGCLFLGNSLRQOQK 120
Db 80 GLQVPEAVNMNLSGLSIAATTSFSQLLAGQPTVHBCGRNMYLTGLCLLQPT--QLTQR 137
Qy 121 FPEALRGCPQEDSDIAFLVDGSGIIPDHFRAKEFISTVMEQLKSKTLFSLMYQSEEF 180
Db 138 LPVSRQECPRQEDIVFLDGGSGISRNPFATMNFVRAVISQFQRPSTQFSLMQFSNKP 197
Qy 181 RIHFTPEFQNNPNRSLKIPITQLLCRTHATGIRKRVRELFTNITNGARKNAFKILILI 240
Db 198 QTHLTPEFRRTNPNLSIASVHQLQGYTTATAIQNVVHLLFHASVGAERDALKILIVI 257
Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFSEKSEKQRIINTVASKPRDHFQIN 300
Db 258 TDGKKEGVDLYKDVIPMAAAGIIRVAGVGLAFQFNENSWKELNDIASKPSQEHIFKVE 317
Qy 301 NFELKTIQNLREKIPAIETGTQSGSSSEHEMSQEGFSAATISNGPLLSITVGSYDAG 360

Db 318 DFDALNDIQOLREKIPPIEGTETTSSTSSFELEMAQEGFSAVFTPDGVLGVSFTWSG 377
Qy 361 GVFLYTSKEKSTFINNRVDSMDNDAYLGYAAAILIIRNRVQSLVGLGAPRYQHILGVAMPR 420
Db 378 GAFLYPNNSPTFINMSQENVDNRDSYLGYSTETLALWKGVSQSLVGLGAPRYQHTKAVIFT 437
Qy 421 QNTGMESANVKGSTOIGAVFGASLCSVDVDSNGSTDVLVIGAPHYHYEQTRGGQSVCP 480
Db 438 QVSRQRMKAETVGTQIGSYFGFSLCSVDVDSNGSTDVLVIGAPHYHYEQTRGQSVCP 497
Qy 481 PRGORAFWQCDAYLGEQGFQWGRFGAALTVLGVDVNGDKLTDVAIGAPGBEDNRGAYL 540
Db 498 PRGWR-RWMCDAVLYGEQGHQWGRFGAALTVLGVDVNGDKLTDVWVGAPGBEENRGAYL 556
Qy 541 HGTSGSISSHSQRISAGLSLPRLOYFGOSLGGDLTWDGLVLTGVAQGHVLLRSO 600
Db 557 HGVLPSPISPSHSQRISAGLSLPRLOYFGOSLGGDLTWDGLVLTGVAQGHVLLRSO 616
Qy 601 PVLRVKXIMEFNRVARNVFECDQVVKGEAGEVVCVLFHVQKSTRDRLEBQIQSV 660
Db 617 PVLWVGYSMOPFIPAEIPRSAFECEQVWSEQLTVQSNCLIVDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAVTNETKNSTRQVGLTQTCETLKLQLPNCIEDPVSPTVLRNF 720
Db 677 LDALDPCGLSPRATFOETKRSLSRVRLGLKAHCENFNLLPSCVEDSVPTITLRNF 736
Qy 721 SLVCTPLSAFQNLPRVLAEDAQRULTALFPEKNCNDNICQDDLSITFSFMSDCLVVG 780
Db 737 TLVGKPLILAFNLPRMLAADAQRYFTASLPEKNCADHICQDNLGISFSFGLKSLVVG 796
Qy 781 GPREFNVTVTVNDGDSYRTQVTFPPDLDSYKRVKSTLQNRQSRWRLACESASSTEV 840
Db 797 SNLEINAEVWVNDGDSYGTITTFPSHPAGLSYRYVAEGQKQQLRSLHLTCDSPAVG-- 854
Qy 841 SGALKSTSCSNHDPFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 855 SQGTWSTSCRINHLIFRGAGITLAFDPSKAVLGRILLITANVSENTPRTSKTTP 914
Qy 901 QLELPVKYAVTVVWVTSQVSTKYLNTAS-ENTSRVMOHQYQVSNLQSRSPISLVFLVP 959
Db 915 QLELPVKYAVTVVWVTSQVSTKYLNTAS-ENTSRVMOHQYQVSNLQSRSPISLVFLVP 974
Qy 960 VRLNQTVINDRPQVTFSENLSTCHTKEELSHSDPLAELKAPVWNCISAVCORIQCDI 1019
Db 975 VELNQAVVMDVSVSLPQNSLCSSEKIAQSPADFLAHIQNPNVLDUSIAGCLAFRCDV 1034
Qy 1020 PFGIQEEFNATLKNLSFDMWIKTSHNELLIVSTAEILFNDSTVTLPGQGAFFVRSOTE 1079
Db 1035 PPSVQOEELDTLKNLSFGVVRQLQKKVSVSVAEITFDTSVYSQLPQGEAFRAQTT 1094
Qy 1080 TKVEPFVNPPLIVGSGVGLLLALITAAIKYLGFPKQYKDMSE 1128
Db 1095 TVLEKVKVHNPTPLIVGSSIGGLLALITAVLYKVGFPKQYKEMMBE 1143
RESULT 13
AAR78166
ID AAR78166 standard; protein; 1161 AA.
XX AAR78166;
XX 28-DEC-1995 (first entry)
XX Human beta-2 integrin alpha-d.
XX Beta-2 integrin alpha-d subunit; antinflammatory; arteriosclerosis;
XX inflammatory bowel disease; asthma.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Domain 17..1108

FT Region /note= "extracellular domain"
 FT 150..352
 FT /note= "this region is homologous to the insert common to
 FT Chla,b,c and may be a site for interaction with ICM
 FT family proteins"
 FT Binding-site 465..474
 FT Binding-site 518..527
 FT Binding-site 592..600
 FT Binding-site 1109..1128
 FT Region /note= "putative cation binding site"
 FT /note= "putative cation binding site"
 FT /note= "transmembrane region"
 FT Domain 1129..1161
 FT /note= "cytoplasmic domain"
 PN W09517412-A1.
 XX
 XX
 PD 29-JUN-1995.
 XX 21-DEC-1994; 94WO-US014832.
 XX 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 XX WPI; 1995-240603/31.
 PI N-PSDB; AAQ91712.
 XX Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 XX Claim 7; Page 82-87; 172pp; English.
 XX A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -TM1 was used to screen a human spleen cDNA library to identify clone
 CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO Cells
 XX Sequence 1161 AA;
 SQ
 Query Match 58.0%; Score 3411; DB 2; Length 1161;
 Best Local Similarity 59.4%; Pred. No. 6.8e-276; Indels 8; Gaps 6;
 Matches 671; Conservative 166; Mismatches 284;
 QY 1 FNLDENAMTFQENARFGQSVVQLQGSRRVVGAPQEIYAANQSGSLYQCYSTGSCPEI 60
 DB 17 FNLDVBEPTIFQEDAGFGQSVVQFGSRLVVGAPLEVVAAANTGRLYDCAAAATGMCPI 76
 QY 61 RLOVFEAVNMSLGLSLAATTSPQILACGPTVHQCSENTYVKGCLFLPSNLRQPOPK 120
 DB 77 PLHIREAVNMSLGLTLAASSTNGSRLACGPTLHRCVGENSYSGKSCLLLSRW-BIIT 135
 QY 121 FPEALRGCPQEDDIAFLVDGSGIIPHDPRRAKEFIETVMEQKKSKTLFSLMOYSEEF 180
 DB 136 VPDATPECHQENDIVFLIDGSGSIDQNDENQMGFQVAVMGQFEGDTLFLALMOYSNLL 195
 QY 181 RHITPKFQNNPNPSLKIPIITOLLGRTHATGIRKYVRELNIINGARKNAFKILLI 240
 DB 196 KIHFTTQFTSPSQSLVDPIVOLKGLTFTATGILFVVTQLFHHKNGARKSAKKILVI 255
 QY 241 TDCKEFGDPLGYEDVPEADREGVIRYVIGWDAPFSEKSRQBLNTVASKPPRDRHVQIN 300
 DB 256 TDCQKYDPLEYSDVPEAKAGIIRYGVGHAFQGTARQELNISSAPQDHFVKYD 315
 QY 301 NPEALKYIQNLRKFIPIAEGTQSGSSFEHMSQEGFSAATISNGPLLTSTVGSVDWAG 360
 DB 316 NFAALGSIQKQLEKIVAVEGTQSRASSSPQHEMSQEGFSTALTMGDLGLGAVGFSWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 376 GAFLYPPNMSPTFINMSQENVDMEDSYLGSTELALWKGVQNLVGLGAPRYQHTGKAVFT 435
 QY 421 QNTGWESNANVKTQTOIGAYFGASLCSDVDVDSNGSTDLVILGAPHYEQTGQGVSVCP 480
 DB 436 QVSRQMKCAEVTOIGSYFGASLCSDVDVDSNGSTDLILIGAPHYEQTGQGVSVCP 495
 QY 481 PRGORARWQCDVILYGEQGPWGRFGAALTIVLGVNDGDKLTVDVAIGAPGEDNRAVYLF 540
 DB 496 PRGORVQWQCDVILYGEQGPWGRFGAALTIVLGVNDGDKLTVDVAIGAPGEDNRAVYLF 555
 QY 541 HGTSGSGLSPSHSORIAGSKLSPLOYFGQSLSGQDLTMDCLVDLTVGAGQGHVLLRSO 600
 DB 556 HGASESGISPSHSORIIASSQLSPLOYFGQSLSGQDLTMDCLVDLTVGAGQGHVLLRSO 615
 QY 601 PVLRVKALMEFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
 DB 616 PVLKVGVMARFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL-GDIQSVR 673
 QY 661 YDLALDSGRPHSRVAVFNETKNTSTRQTVLGLTOTCETLKLQPNCTIEDPVSIVLRNLF 720
 DB 674 FDLALDPGLRTSRAIFNETKNTSTRQTVLGLTOTCETLKLQPNCTIEDPVSIVLRNLF 733
 QY 721 SLVGTPLSAFCNLRPVLAEADAQRELTALFPPEKNCGMNDIQQDDLSITFSFMSLDCLVYG 780
 DB 734 SLVREPSPQNLAPVLAVGQDLFTASLPPEKNGQDGLCEGDLGVTLSFSGLTLTWG 793
 QY 781 GPRFNVTVTVRNDGDSYRTQVTFPPPLDLSYKVKVSTLQNRQSRWRLACESASSTEV 840
 DB 794 SSLELNVTVTVMNAGEDSYGTWVSLYTPAGLSHRVSGAQPHQSALRLACETV-PTED 852
 QY 841 SGALKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPDNKTEF 900
 DB 853 EG-LRSSKCSVNHPIFHEGSGTIFIVFDVSKATLGDRLMARASSENKASSSKATP 911
 QY 901 QLELPKVAVTVMVTSHGVTGYLNF-TASENTRVMQHOYQVGNLNGORSIPISLVFLVP 959
 DB 912 QLELPKVAVTVMISRQESTKYFNFATSDEKKKKEAHRVYRVNLSQORDLAISINFWVP 971
 QY 960 VRLNQTVINDRPQVTFPSNLSSTCTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDI 1019
 DB 972 VLLNGVAVDVMRAPSQSL--PCVSEKRPQHSDFLTQISRPMLDCSIADCLQFRCDV 1029
 QY 1020 PFFGIQEEFNATLKNLSFDWYIKTISENHLILVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
 DB 1030 PSFSVQSELDFTLKNLSFGWVRETLQKKVLVSVVAEITPDTSVVSQLPGQCAFRAQME 1089
 QY 1080 TKVPEPFVNPPLIVGSSVCGLLILLALITAAALYKLGFFKRYKMKSE 1128
 DB 1090 MVLEDEYVNAIPTIMGSSVGALLLALITATLYKLGFFKRYKMKSE 1138
 RESULT 14
 AAW23049
 ID AAW23049 standard; protein; 1161 AA.
 XX
 AC AAW23049;
 XX
 DT 24-FEB-1998 (first entry)
 XX
 DE Human beta 2 integrin alpha d subunit.
 XX
 XX Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW psoriasis; lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Domain 17..1108
 FT

```

PT      /label= Extracellular_domain
FT      150..352
FT      /note= "region homologous to the I (insertion) domain
FT      common to CD11a, CD11b and CD11c"
FT      1109..1128
FT      /label= Transmembrane domain
FT      /note= "homologous to the human CD11c transmembrane
FT      region"
FT      1129..1161
FT      /label= Cytoplasmic_domain
XX      WO9731099-A1.
XX      28-AUG-1997.
XX      24-FEB-1997; 97WO-US002713.
XX      22-FEB-1996; 96US-00605672.
XX      (ICOS-) ICOS CORP.
XX      PA
XX      Gallatin WM, Van Der Vieren M;
XX      WPI; 1997-435154/40.
XX      DR      N-PSDB; AAT79220.
XX      DR
XX      Hybridoma 199M and antibody secreted by it - specific for new rat beta2
XX      integrin subunit, useful to detect subunit in cells and modulate its
XX      activity.
XX      Example 5; Page 116-120; 222pp; English.
XX      This polypeptide comprises a novel human beta 2 integrin subunit,
XX      designated alpha d. Its sequence was deduced from a cDNA clone (see
XX      AAT79220) isolated from a spleen cDNA library. Alpha d is involved in
XX      cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
XX      d polypeptides can be expressed in transformed host cells for use in
XX      assays for identifying antibodies or other compounds that modulate alpha
XX      d activity or which modulate the interaction between alpha d and a
XX      ligand, for treating or preventing diseases in which macrophages are
XX      implicated. Treatment is applicable to disease states in which alpha d
XX      binding or localised accumulation of cells which express alpha d, is
XX      implicated such as such as type I diabetes, atherosclerosis, multiple
XX      sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX      distress syndrome and rheumatoid arthritis
XX      SQ      Sequence 1161 AA;
XX      Query Match 58.0%; Score 3411; DB 2; Length 1161;
XX      Best Local Similarity 59.4%; Pred. No. 6.8e-276;
XX      Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
XX      1 FNLDTNMTFOENARGFGQSVVQLOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPTI 60
XX      17 FNLDVREPTFQEDAGFGQSVVQFGSRLVVGAPLEVVAAANTGRLYDCAATGMCPTI 76
XX      61 RLQVPEAVNMSGLSLAATTSPPQLACGPTVHTQCSNTYVKGCCFLGSLNRQPOK 120
XX      77 PLHIREAVNMSGLTAASTNGSLRACGPTLHRVCGENSYKSGCLLIGSRW-ELIQT 135
XX      121 FPBALRGCPQEDSDIAPLVDSGSIIPHDFERRAKEFISTVMEQLKSKTILFSLMOYSBEF 180
XX      136 VPDATPCPHQEMDIVPLIDSGSIDQNDPNQMGFVQAVMGQFEGDTLIFALMOYSNEL 195
XX      181 RIHPTFEKQNNPNSRLIKPIQLIGRTHATGIRKVVRELNIYNGARKNAFKILLI 240
XX      196 KIHFTPTQFTSPSQSLVDPIVLQKLTETATGILTVTQLFHHKNGARKSAKILLI 255
XX      241 TDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRSLNTVASPPRDHVFQIN 300
XX      256 TDGQKYKDPLEYSNVIQAEKAGIIRVAIGVGHAFQGPPIARQELNISSAPPDHFVKVD 315
XX      301 NFEALKTIQNLREKIFAIETGTGTGSSSPSEHEMSQSGFSAATISNGPLISTVGSYDWAG 360

```

```

Db      316 NFAALGSIQKQLQEKI VAVEGTQSRASSSPQHEMSQSGFSTALTWGDLFLGAVGSFMSG 375
Qy      361 GVFLYTSKEKSTFINMTRVSDMDAYLGYAAAIILNRVQSIVLFGARYOHIGIVAMER 420
Db      376 GAFLYPPNMSFTFINMSQENVDMDSYLGYSSTELALWGVQNLVLGARYOHTGKAVIPT 435
Qy      421 QNTGWTENANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVSCPL 480
Db      436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVSCPL 495
Qy      491 PRGQARWQCDAYLYGBOGQPWGFGAALTVLGVNGDKLTDVAIGAPCEEDNGAVLYF 540
Db      496 PRGQRVQWQCDAYLYGBOGQPWGFGAALTVLGVNGDKLTDVAIGAPCEEDNGAVLYF 555
Qy      541 HGTSGGISPSHSORIKSILSPRLQYFGQSLGSGQDLTMDGLVLDLTVGAQCHVLLRSQ 600
Db      556 HGASEGISPSHSORIKSILSPRLQYFGQSLGSGQDLTMDGLVLDLTVGAQCHVLLRSQ 615
Qy      601 PVLRYKALMEFNPREVARNFECDNDQVVKGEAGBVRVCLHVQKSTRDLREGQIQSVVT 660
Db      616 PVLKVGAMRFPVEVAKAVRCWEKEPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
Qy      661 YDLALDSGRPHSRVAFNETKSTRTOVLGLTTCETLKLQLPNCIEDPVPVILRLNF 720
Db      674 FDLALDPGRLSRAIFNETKPTLTRKYLGLGHCETLKLKLLPCVEDVVSPIILHLNF 733
Qy      721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALFPPEKKGNDNICODDLSITFSFMSLCLVVG 780
Db      734 SLVREPIPSQNLRFVLAAGSQDLFTASLPPEKKGQGLCEGDLGVTLSFSGLTLLTVG 793
Qy      781 GPREENVTIVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNGRSQSWRLACESASSTEV 840
Db      794 SSLELNVTIVVWNGEDSIGTVSDIYPAGLSHRVSGAQKQPHQSALRLACETV-PTED 852
Qy      841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
Db      853 EG-LSSRCVSNHPIFEGSGNGTFIVTFDVSYKATLGDRLMRLRAGASSENKASSKATF 911
Qy      901 QLELPVKIVAMVWTSVSHGVSKYLNFTASENTSVMOHVOYVNSLQORSPLSLVFLVP 959
Db      912 QLELPVKIVAMVWTSVSHGVSKYLNFTASENTSVMOHVOYVNSLQORSPLSLVFLVP 971
Qy      960 VELNQTIVNDRPQVTFSENLSSTCHTKERLPSHSDFLAELAKAPVNVGSIACQRIQCDI 1019
Db      972 VLLNGVAVDWVMEAPSQSL--PCVSEKPKQHSDFLTQISPSMLDCSIADCLQPCDV 1029
Qy      1020 PFGIOREFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVTLTLPQGQAFVRSQTE 1079
Db      1030 PFSVQOEELDFTLKGNLSFGWVRETLOKKVLVSVVAEITFDTSVYSQLPGQAFVRAQME 1089
Qy      1080 TKVEFPFVNPPLIPLVSGSVGLLALLITLITLALYKLGFFKQYKDWMS 1128
Db      1090 MVLEDEVTNALPIITGSSVGLLALLITLITLALYKLGFFKQYKDWMS 1138
XX      RESULT 15
XX      AAW57491
XX      ID      AAW57491 standard; protein; 1161 AA.
XX      AC      AAW57491;
XX      XX
XX      DT      24-AUG-1998 (first entry)
XX      XX
XX      DE      Human beta2 integrin alpha subunit (alpha d) polypeptide.
XX      KW      Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;
XX      reporter-transactivator construct; arteriosclerosis; atherosclerosis;
XX      inflammatory bowel disease; arthritis; multiple sclerosis.
XX      OS      Homo sapiens.
XX      PH      Key
XX      Location/Qualifiers

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds
(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKQYKDMSEGGPGGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5876	100.0	1137	10	US-09-902-481A-5
2	5862	99.8	1153	9	US-09-350-259-3
3	5862	99.8	1153	10	US-09-902-481A-1
4	5862	99.8	1153	14	US-09-891-943-3
5	5862	99.8	1153	14	US-10-144-259-30
6	5862	99.8	1153	14	US-10-207-655-176
7	5855	99.6	1137	10	US-09-902-481A-6
8	5851	99.6	1137	10	US-09-902-481A-4
9	5846.5	99.5	1152	9	US-09-945-265-4
10	5836	99.3	1137	10	US-09-902-481A-3
11	3469	59.0	1163	15	US-10-116-275-204
12	3446	58.6	1163	9	US-09-350-259-4
13	3446	58.6	1163	9	US-09-891-943-4
14	3411	58.0	1161	9	US-09-350-259-2
15	3411	58.0	1161	10	US-09-891-943-2

16	3395.5	57.8	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3395.5	57.8	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3332.5	55.0	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3332.5	55.0	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3222	54.8	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3222	54.8	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3211.5	54.7	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3211.5	54.7	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3201	54.5	1151	9	US-09-350-259-46	Sequence 46, Appl
25	3201	54.5	1151	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	12	US-10-087-192-1212	Sequence 1212, Ap
27	1557.5	26.5	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1553.5	26.4	1170	12	US-10-261-164-1	Sequence 1, Appli
29	1530	26.0	1223	16	US-10-408-765A-295	Sequence 295, App
30	1359.5	23.1	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1153	19.6	1179	14	US-10-177-550-2	Sequence 2, Appli
34	1153	19.6	1179	15	US-10-173-551-2	Sequence 101, App
35	1151.5	19.6	413	9	US-09-350-259-101	Sequence 101, App
36	1151.5	19.6	413	10	US-09-891-943-101	Sequence 101, App
37	1103.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
38	1103.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
39	1103.5	18.8	1179	12	US-09-918-715-250	Sequence 250, App
40	1101	18.7	1188	15	US-10-291-265-810	Sequence 810, App
41	1097	18.7	1188	15	US-10-291-265-338	Sequence 338, App
42	1093.5	18.6	1189	10	US-09-984-130-35	Sequence 35, Appl
43	1093.5	18.6	1189	10	US-09-836-353A-35	Sequence 35, Appl
44	1093.5	18.6	1189	12	US-10-262-839-4	Sequence 4, Appli
45	1084	18.4	589	12	US-10-261-164-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-902-481A-5
; Sequence 5, Application US/0902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match	100.0%;	Score 5876;	DB 10;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQPIVANQSGLSYQCDYSTGSCPEI	60	
Db	1	FNLDTENAMTFQENARGFGQSVVOLQSGRVVVGAPQPIVANQSGLSYQCDYSTGSCPEI	60	
Qy	61	RLOVPVAVNMSLGLSLAATTPQLACGPTVHTCSTENTYVKGCLCFPGSNLRQOPQK	120	
Db	61	RLOVPVAVNMSLGLSLAATTPQLACGPTVHTCSTENTYVKGCLCFPGSNLRQOPQK	120	
Qy	121	PFREALGCPQEDSDIAFLVDGSGSIIPHDFRAKEFTSTVMEQLKSKTLFSLMOYSEEP	180	

Db	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQKKSKTLFSLMQVSEFP	180
Qy	181	RIHFTPEFQNNPNSRLKPIITQLGRTHATGIRKVVRELFNTNGARKNAFKLILILI	240
Db	181	RIHFTPEFQNNPNSRLKPIITQLGRTHATGIRKVVRELFNTNGARKNAFKLILILI	240
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNVTASKPPRDHVFQIN	300
Db	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNVTASKPPRDHVFQIN	300
Qy	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAQ	360
Db	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAQ	360
Qy	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR	420
Db	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR	420
Qy	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL	480
Db	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL	480
Qy	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Db	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Qy	541	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSCQDPLTMDGLVDLTGVAQGHVLLRSQ	600
Db	541	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSCQDPLTMDGLVDLTGVAQGHVLLRSQ	600
Qy	601	PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Db	601	PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Qy	661	YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720
Db	661	YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720
Qy	721	SLVGTPLSAFQNLAPVLAEDAQRULTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG	780
Db	721	SLVGTPLSAFQNLAPVLAEDAQRULTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG	780
Qy	781	GPREFNVTVVRNDGDSYRQVTFPPPLDLISYKVVSTLQNRQSQRWRLACSSASTEV	840
Db	781	GPREFNVTVVRNDGDSYRQVTFPPPLDLISYKVVSTLQNRQSQRWRLACSSASTEV	840
Qy	841	SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKXANTSENMTPTNKTEF	900
Db	841	SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKXANTSENMTPTNKTEF	900
Qy	901	QLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV	960
Db	901	QLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV	960
Qy	961	RINQTVIWDPRQVTFENSELSTCHTKERLPSHSDFLAELRKA P VVNCISIAVCQRIQDIP	1020
Db	961	RINQTVIWDPRQVTFENSELSTCHTKERLPSHSDFLAELRKA P VVNCISIAVCQRIQDIP	1020
Qy	1021	PFQIQEENATLKNLSFDWIKTSHNHLIIVSTABEILFNDSTVFTLLPGQAGFVRSQTET	1080
Db	1021	PFQIQEENATLKNLSFDWIKTSHNHLIIVSTABEILFNDSTVFTLLPGQAGFVRSQTET	1080
Qy	1081	KVEPFEPVNPPLIVGSGVGLLILALITAAALYKLGFPFKQYKDMXSEGGPPGAEPPQ	1137
Db	1081	KVEPFEPVNPPLIVGSGVGLLILALITAAALYKLGFPFKQYKDMXSEGGPPGAEPPQ	1137

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US2002062008A1
; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2002062008A1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/350,259
CURRENT FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: 09/193,043
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
US-09-350-259-3

Query Match	99.8%;	Score 5862;	DB 9;	Length 1153;
Best Local Similarity	99.3%;	Pred. No. 0;		
Matches 1129;	Conservative	7;	Mismatches	1; Indels 0; Gaps 0;
QY	1	FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPOEIVAAHQRSGLYQCDYSTGSCPEI	60	
Db	17	FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPOEIVAAHQRSGLYQCDYSTGSCPEI	76	
QY	61	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPGSNLRQPOPK	120	
Db	77	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPGSNLRQPOPK	136	
QY	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQLKSKTLFSLMQVSEEF	180	
Db	137	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTSTVMEQLKSKTLFSLMQVSEEF	196	
QY	181	RIHFTPEKFOFNNPNSRLKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKLILILI	240	
Db	197	RIHFTPEKFOFNNPNSRLKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKLILIVI	256	
QY	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNVTASKPPRDHVFQIN	300	
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNVTASKPPRDHVFQIN	316	
QY	301	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAQ	360	
Db	317	NFEALKTIONOLREKIPALTEGTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAQ	376	
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR	420	
Db	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR	436	
QY	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL	480	
Db	437	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL	496	
QY	481	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	540	
Db	497	PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF	556	
QY	541	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSCGQDPLTMDGLVDLTGVAQGHVLLRSQ	600	
Db	557	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSCGQDPLTMDGLVDLTGVAQGHVLLRSQ	616	
QY	601	PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660	
Db	617	PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	676	
QY	661	YDLALDSGRPHSAVFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720	

```

Db      677 YDLDLDSGRPHSAVFNENKSTRTQVGLGLTQTCETLKLQPLNCIEDPVSIVLRLNF 736
Qy      721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKMGNDNICODDLSITFSFMSLDCLVVG 780
Db      737 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKMGNDNICODDLSITFSFMSLDCLVVG 796
Qy      781 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKVSITLQNSQSRSLACSSASTEV 840
Db      797 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKVSITLQNSQSRSLACSSASTEV 856
Qy      841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTPE 900
Db      857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTPE 916
Qy      901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHYQVSNLQGRSLPISLVFLVPV 960
Db      917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHYQVSNLQGRSLPISLVFLVPV 976
Qy      961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVVMCSIAVCORIOCDIP 1020
Db      977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVVMCSIAVCORIOCDIP 1036
Qy      1021 FFGIOEFNATLKGNSLDFDWTIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTEP 1080
Db      1037 FFGIOEFNATLKGNSLDFDWTIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTEP 1096
Qy      1081 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1137
Db      1097 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1153

```

RESULT 3

```

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

```

```

Query Match      99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy      1  FNLDTENAMTFOENARGFGQSVVQGSRRVVGAPQIVAAQNSLYQDYSTGSCBPI 60
Db      17  FNLDTENAMTFOENARGFGQSVVQGSRRVVGAPQIVAAQNSLYQDYSTGSCBPI 76

Qy      61  RLQVPEAVNMSGLSLAATSPKLLACGPTTHQTCSENTYKGLCFPGSNLRQOPK 120
Db      77  RLQVPEAVNMSGLSLAATSPKLLACGPTTHQTCSENTYKGLCFPGSNLRQOPK 136

Qy      121 FPEALRGCPQSDSIAFLIDGSGSIIIPHDFFRMKEFVSTVMEQLKSKTLFSLMOYSEEF 196

```

```

Db      137 FPEALRGCPQSDSIAFLIDGSGSIIIPHDFFRMKEFVSTVMEQLKSKTLFSLMOYSEEF 196
Qy      181 RIHFTFKFQNNPNSRSLIKPITOLLGRTHTATGIRKVVRELFNITNGARNAKILLILI 240
Db      197 RIHFTFKFQNNPNSRSLIKPITOLLGRTHTATGIRKVVRELFNITNGARNAKILLILI 256
Qy      241 TDGKFGDPLGYEYDVIPEADREGVIRYVIGVGDAPFRSEKSRQELMTASKPPRDHVFQIN 300
Db      257 TDGKFGDPLGYEYDVIPEADREGVIRYVIGVGDAPFRSEKSRQELMTASKPPRDHVFQIN 316
Qy      301 NPEALKTIQNLREKIPAIETCTOTGSSSSFEHMSORSPSAATITNSGPLLSTVGSVDWAG 360
Db      317 NPEALKTIQNLREKIPAIETCTOTGSSSSFEHMSORSPSAATITNSGPLLSTVGSVDWAG 376
Qy      361 GVFLYTSSEKSTFNMTRVDSMDNDAYLVAAAIILNRVQSVILGAPRYOHICLVAMFR 420
Db      377 GVFLYTSSEKSTFNMTRVDSMDNDAYLVAAAIILNRVQSVILGAPRYOHICLVAMFR 436
Qy      421 QNTGMBESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQRGGQSVCP 480
Db      437 QNTGMBESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQRGGQSVCP 496
Qy      481 PRGORARWQCDVILYGRGQGWGRFGAALTVLGDVNGDKLTDVAIGAPGBEDNRGAYL 540
Db      497 PRGORARWQCDVILYGRGQGWGRFGAALTVLGDVNGDKLTDVAIGAPGBEDNRGAYL 556
Qy      541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTWDGELVDTLVGAQGHVLLRSQ 600
Db      557 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTWDGELVDTLVGAQGHVLLRSQ 616
Qy      601 PVLRYKAIMFNPREVARNVFECDQVYVKGKAGEVRLVHVKSTDRDLRBSQIQSVVT 660
Db      617 PVLRYKAIMFNPREVARNVFECDQVYVKGKAGEVRLVHVKSTDRDLRBSQIQSVVT 676
Qy      661 YDLALDSGRPHSAVFNENKSTRTQVGLGLTQTCETLKLQPLNCIEDPVSIVLRLNF 720
Db      677 YDLALDSGRPHSAVFNENKSTRTQVGLGLTQTCETLKLQPLNCIEDPVSIVLRLNF 736
Qy      721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKMGNDNICODDLSITFSFMSLDCLVVG 780
Db      737 SLVGTPLSAFGLNRPVLAEDAQRFTALPPEKMGNDNICODDLSITFSFMSLDCLVVG 796
Qy      781 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKVSITLQNSQSRSLACSSASTEV 840
Db      797 GPRFNVTVTVRNDEGDSYRTQVTFPPDLQSVKVSITLQNSQSRSLACSSASTEV 856
Qy      841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTPE 900
Db      857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTPE 916
Qy      901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHYQVSNLQGRSLPISLVFLVPV 960
Db      917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHYQVSNLQGRSLPISLVFLVPV 976
Qy      961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVVMCSIAVCORIOCDIP 1020
Db      977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVVMCSIAVCORIOCDIP 1036
Qy      1021 FFGIOEFNATLKGNSLDFDWTIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTEP 1080
Db      1037 FFGIOEFNATLKGNSLDFDWTIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTEP 1096
Qy      1081 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1137
Db      1097 KVPEPEVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKXMMSEGGPPGAEPO 1153

```

RESULT 4

```

US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US2003007728A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.

```

APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US20030077278A1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI 76

QY 61 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 61 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRAMEPISITVMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRAMEPISITVMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 256

APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US20030077278A1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI 76

QY 61 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFLFGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRAMEPISITVMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRAMEPISITVMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNPSLKPITQLLGRTHATGIRKVVRELFTNNGARKNAFKILILI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQBLNTVASKPRDHVQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQBLNTVASKPRDHVQIN 316

QY 301 NFEALKTIQNLREKIPAIETGTQSSSSSEHEMSQSGFSAATNSGPLLSTVGSYDMAG 360
DB 317 NFEALKTIQNLREKIPAIETGTQSSSSSEHEMSQSGFSAATNSGPLLSTVGSYDMAG 376

QY 361 GVFLYTSKSKSTFNTMRVDSMDMDAYLGAAAILLNRVQSLVGLGAPRYCHIGLVAMPR 420
DB 377 GVFLYTSKSKSTFNTMRVDSMDMDAYLGAAAILLNRVQSLVGLGAPRYCHIGLVAMPR 436

QY 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL 480
DB 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL 496

QY 481 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYTLF 540
DB 497 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYTLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTWDGLVDLTIVGAGQHVILLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTWDGLVDLTIVGAGQHVILLRSQ 616

QY 601 PVLKALMEPNPREVARNVECDQVVKGEAGEVAVLHVQKSTDRLRBEGQIQSVVT 660
DB 617 PVLKALMEPNPREVARNVECDQVVKGEAGEVAVLHVQKSTDRLRBEGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNFTKNSRQTOVLGLTQTCETFLKQLQNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAFNFTKNSRQTOVLGLTQTCETFLKQLQNCIEDPVPVILRLNF 736

QY 241 TDEKEFGDPLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDRHVQIN 300
DB 257 TDEKEFGDPLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDRHVQIN 316
QY 301 NFEALXATIQOLREKIFAIEGTOTGSSSFEHMSQEGFSAATTSNGPLLSVGSVDWAG 360
DB 317 NFEALXATIQOLREKIFAIEGTOTGSSSFEHMSQEGFSAATTSNGPLLSVGSVDWAG 376
QY 361 GVELYTSKEKSTPINTRVDSMDNDAYLGVAALIIIRNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVELYTSKEKSTPINTRVDSMDNDAYLGVAALIIIRNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEGQOPWRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORARWQCDVLYGEGQOPWRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPCIEDPVSIVLRNLF 720
DB 677 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPCIEDPVSIVLRNLF 736
QY 721 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVTVSHGVSSTKYLNFNTASENTSRVMOHQVQVSNLQORSPLISLVELVPV 960
DB 917 QLELPVKYAVYVTVSHGVSSTKYLNFNTASENTSRVMOHQVQVSNLQORSPLISLVELVPV 976
QY 961 RLNQTWINDRPOVTSSENLSSTCHTERLPSSHDFLAEIRKAPVWNCSTAVCORIQCDIP 1020
DB 977 RLNQTWINDRPOVTSSENLSSTCHTERLPSSHDFLAEIRKAPVWNCSTAVCORIQCDIP 1036
QY 1021 PFGIQEENATLKGNSLSPWYIKNTSHNLLIYSTABILEFNDVSFTLLPQCGAFVRSQTE 1080
DB 1037 PFGIQEENATLKGNSLSPWYIKNTSHNLLIYSTABILEFNDVSFTLLPQCGAFVRSQTE 1096
QY 1081 KVEPPFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFRQYKDMKSEGPPGABEPQ 1137
DB 1097 KVEPPFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFRQYKDMKSEGPPGABEPQ 1153

RESULT 6
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.8%; Score 5862; DB 14; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQBEIVAAANQORSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQBEIVAAANQORSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHGTCSNTYVYKGLCLFPGSLNRQQPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHGTCSNTYVYKGLCLFPGSLNRQQPQK 136
QY 121 PPEARLGGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIITVMEQLKSKTLPFLSMOYSSEF 180
DB 137 PPEARLGGCPQEDSDIAFLIDSGSGSIIPHDPRRAKEFIITVMEQLKSKTLPFLSMOYSSEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATATGIRKVVRELFTNIGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATATGIRKVVRELFTNIGARKNAFKILIVI 256
QY 241 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDRHVQIN 300
DB 257 TDGEKFGPGLGYEDVTPADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDRHVQIN 316
QY 301 NFEALXATIQOLREKIFAIEGTOTGSSSFEHMSQEGFSAATTSNGPLLSVGSVDWAG 360
DB 317 NFEALXATIQOLREKIFAIEGTOTGSSSFEHMSQEGFSAATTSNGPLLSVGSVDWAG 376
QY 361 GVELYTSKEKSTPINTRVDSMDNDAYLGVAALIIIRNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVELYTSKEKSTPINTRVDSMDNDAYLGVAALIIIRNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEGQOPWRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORARWQCDVLYGEGQOPWRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGOSISGGQDLTMDGLVDLTGAGQHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPCIEDPVSIVLRNLF 720
DB 677 YDLALDSGRPHSRAVFNETHKSTRTOTVGLTQTCETLKLQPCIEDPVSIVLRNLF 736
QY 721 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEADQRLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDCGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVTVSHGVSSTKYLNFNTASENTSRVMOHQVQVSNLQORSPLISLVELVPV 960

Db 917 QLELPVKYAVVMTSHGVTSTKILNFTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVINDRPQVTFSENLSSSTCHTKERLPSSDPLAELRKAPVNVNCISIAVCQRIQCDDIP 1020
Db 977 RLNQTVINDRPQVTFSENLSSSTCHTKERLPSSDPLAELRKAPVNVNCISIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEIILFNDVFTLLPQGGAFVRSQET 1080
Db 1037 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEIILFNDVFTLLPQGGAFVRSQET 1096
Qy 1081 KVEPPEVNPFLPLIVGSSVGGILLALITAAALYKLGFPKQYKQKMMWSEGGPPGABPQ 1137
Db 1097 KVEPPEVNPFLPLIVGSSVGGILLALITAAALYKLGFPKQYKQKMMWSEGGPPGABPQ 1153

RESULT 7
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.6%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQNRAGFGQSVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCDYSTGSCBEI 60
Db 1 FNLDTENAMTFQNRAGFGQSVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCDYSTGSCBEI 60
Qy 61 RLQVPEAVNMVSLGLSAAITTSPPQLACGPTVHQTCSENTYVYKGLCFGLGSLNRQOPQK 120
Db 61 RLQVPEAVNMVSLGLSAAITTSPPQLACGPTVHQTCSENTYVYKGLCFGLGSLNRQOPQK 120
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRAXEFISTVMEQLKSKTILFSLMOYSEEP 180
Db 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRAXEFISTVMEQLKSKTILFSLMOYSEEP 180
Qy 181 RIHFTKSPQNNPNSLKIPIITOLGRTHATGIRKVVRELVNITNGAKNAFKILILI 240
Db 181 RIHFTKSPQNNPNSLKIPIITOLGRTHATGIRKVVRELVNITNGAKNAFKILILI 240
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVIGVDGDAFRSEKSRQELNTVASKPPRHDVFPQIN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRVYVIGVDGDAFRSEKSRQELNTVASKPPRHDVFPQIN 300
Qy 301 NFPAKTIQNLREKIFAIGTQTSSEHMSQEGFSAATISNGPLISTVGSYDMAG 360
Db 301 NFPAKTIQNLREKIFAIGTQTSSEHMSQEGFSAATISNGPLISTVGSYDMAG 360
Qy 361 GVFLYTSKEKSTFNMTRVDSMDNDAYLGVAAILLNRRVQSLVGLAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFNMTRVDSMDNDAYLGVAAILLNRRVQSLVGLAPRYQHIGLVAMFR 420
Qy 421 QNTGMWESNANVKQTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCPCL 480

Db 421 QNTGMWESNANVKQTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCPCL 480
Qy 481 PRGORARWQCDALVYGBQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNEGAVYLF 540
Db 481 PRGORARWQCDALVYGBQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNEGAVYLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTGAGRHVLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTGAGRHVLLRSQ 600
Qy 601 PVLRVKAIMFENPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
Db 601 PVLRVKAIMFENPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
Qy 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTCTETLKLQLPNCIEDPVSPIVLRINF 720
Db 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTCTETLKLQLPNCIEDPVSPIVLRINF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPEKNCNGDNICQDDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPEKNCNGDNICQDDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQORSWRWLACESASSTEV 840
Db 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRQORSWRWLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENKMPRTNKTEP 900
Db 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENKMPRTNKTEP 900
Qy 901 QLELPVKYAVVMTSHGVTSTKILNFTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 901 QLELPVKYAVVMTSHGVTSTKILNFTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Qy 961 RLNQTVINDRPQVTFSENLSSSTCHTKERLPSSDPLAELRKAPVNVNCISIAVCQRIQCDDIP 1020
Db 961 RLNQTVINDRPQVTFSENLSSSTCHTKERLPSSDPLAELRKAPVNVNCISIAVCQRIQCDDIP 1020
Qy 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEIILFNDVFTLLPQGGAFVRSQET 1080
Db 1021 PFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEIILFNDVFTLLPQGGAFVRSQET 1080
Qy 1081 KVEPPEVNPFLPLIVGSSVGGILLALITAAALYKLGFPKQYKQKMMWSEGGPPGABPQ 1137
Db 1081 KVEPPEVNPFLPLIVGSSVGGILLALITAAALYKLGFPKQYKQKMMWSEGGPPGABPQ 1137

RESULT 8

US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 99.6%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFOENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

Qy 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120

Qy 121 FPEALRGCPQSDSDIAFLVDSGSGIIIPHDPRAKEFTSTVMEQKSKTLFSLMQYSEEF 180
Db 121 FPEALRGCPQSDSDIAFLVDSGSGIIIPHDPRAKEFTSTVMEQKSKTLFSLMQYSEEF 180

Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240

Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 300
Db 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 300

Qy 301 NFEALKTIONLREKI PAIEGTQTGSSSSFEHMSQSGFSAITNSNGSTDLVIGAPHYEOTRGQSVCP 360
Db 301 NFEALKTIONLREKI PAIEGTQTGSSSSFEHMSQSGFSAITNSNGSTDLVIGAPHYEOTRGQSVCP 360

Qy 361 GVFLYTSKEKSTFNTMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFNTMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQSVCP 480
Db 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQSVCP 480

Qy 481 PRGQARWQCDVLYGEGQPKWGFALITVLGVNGDKLTDVAIGAPGEDNEGAVLYF 540
Db 481 PRGQARWQCDVLYGEGQPKWGFALITVLGVNGDKLTDVAIGAPGEDNEGAVLYF 540

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTAVGAGHYLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTAVGAGHYLLRSQ 600

Qy 601 PVLAVKALMEFPREAVRNFECDQVVKGEAEVRLVHVKSTRDRLEGOIQSVYT 660
Db 601 PVLAVKALMEFPREAVRNFECDQVVKGEAEVRLVHVKSTRDRLEGOIQSVYT 660

Qy 661 YDLALDSGRPHSRVAFNETNSTRTQVGLTQTCETLKLQLENCIEDPVPVILRLNF 720
Db 661 YDLALDSGRPHSRVAFNETNSTRTQVGLTQTCETLKLQLENCIEDPVPVILRLNF 720

Qy 721 SLVGTPLSAGNLAPVLAEDAQRLPTALFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAGNLAPVLAEDAQRLPTALFPEKNCNDNICODDLSITFSFMSLDCLVVG 780

Qy 781 GPREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 781 GPREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840

Qy 841 SGALKSTSCSINHPIPEPNSVNTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 841 SGALKSTSCSINHPIPEPNSVNTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900

Qy 901 QLELPVKYAVTMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSILPISLVLFPV 960
Db 901 QLELPVKYAVTMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSILPISLVLFPV 960

Qy 961 RLNQTIVDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVWNCIAVCQRIQCDIP 1020
Db 961 RLNQTIVDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVWNCIAVCQRIQCDIP 1020

Qy 1021 FFGIQEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080

Db 1021 FFGIQEFNATLKGNSLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080

Qy 1081 KVEPFEPVNPPLIIVGSSVGGILLALLIITAAALVKLGFPRKQYKDMSEGGPPGAEPO 1137
Db 1081 KVEPFEPVNPPLIIVGSSVGGILLALLIITAAALVKLGFPRKQYKDMSEGGPPGAEPO 1137

RESULT 9
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimooka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CHN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.5%; Score 5846.5; DB 9; Length 1152;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDSDIAFLVDSGSGIIIPHDPRAKEFTSTVMEQKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLVDSGSGIIIPHDPRAKEFTSTVMEQKSKTLFSLMQYSEEF 196

Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 316

Qy 301 NFEALKTIONLREKI PAIEGTQTGSSSSFEHMSQSGFSAITNSNGSTDLVIGAPHYEOTRGQSVCP 360
Db 317 NFEALKTIONLREKI PAIEGTQTGSSSSFEHMSQSGFSAITNSNGSTDLVIGAPHYEOTRGQSVCP 376

Qy 361 GVFLYTSKEKSTFNTMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFNTMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQSVCP 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQSVCP 496

Qy 481 PRGQARWQCDVLYGEGQPKWGFALITVLGVNGDKLTDVAIGAPGEDNEGAVLYF 540
Db 497 PRGQARWQCDVLYGEGQPKWGFALITVLGVNGDKLTDVAIGAPGEDNEGAVLYF 555

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTAVGAGHYLLRSQ 600

DB 556 HGTSGSISPSHSQRIAGSKLSPLQVFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 615
QY 601 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLFVQKSTRDLREGQIQSVVT 660
DB 616 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLFVQKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 676 YDLALDSGRPHSRVAFNETKSTRQTVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
QY 721 SLVGTPLSAFQNLAPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFQNLAPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLVSVEKVSFELQNSQSRWRLACESASTEV 840
DB 796 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLVSVEKVSFELQNSQSRWRLACESASTEV 855
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
DB 856 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 915
QY 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTSRVMQHQYQVSNLQORSILPLVPLVPV 960
DB 916 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTSRVMQHQYQVSNLQORSILPLVPLVPV 975
QY 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1020
DB 976 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1035
QY 1021 PFGIQEENATLKGNSLDFWYIKTSHNHLLIVSTAEILFNDVSFTLLPQAGAFVRSQTEF 1080
DB 1036 PFGIQEENATLKGNSLDFWYIKTSHNHLLIVSTAEILFNDVSFTLLPQAGAFVRSQTEF 1095
QY 1081 KVEPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFPKQYKQKDMWSEGGPPGABPQ 1137
DB 1096 KVEPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFPKQYKQKDMWSEGGPPGABPQ 1152

RESULT 10

US-09-902-481A-3

; Sequence 3: Application US/0902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimooka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/502,481A

; CURRENT FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,600

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-3

Query Match 99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQBIIVAAHQSLVQCDYSTGSCBPI 60
DB 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQBIIVAAHQSLVQCDYSTGSCBPI 60
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLFLFGLNLRQOPQK 120

RESULT 11

DB 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLFLFGLNLRQOPQK 120
QY 121 FPEALRGCPQSDSIAELVDGSGSIIIPHDERRAKEFTSTVBOELKKSKTILFSLMOYSEBF 180
DB 121 FPEALRGCPQSDSIAELVDGSGSIIIPHDERRAKEFTSTVBOELKKSKTILFSLMOYSEBF 180
QY 181 RIHTTFKEFQNNPRLSLIKPIITOLLORTHTATGIRKQVRELFNITNGARKNAFKILILI 240
DB 181 RIHTTFKEFQNNPRLSLIKPIITOLLORTHTATGIRKQVRELFNITNGARKNAFKILILI 240
QY 241 TDGEKFGDPLGYEDVIVPEADREGVIRYVIGVDAFRSEKSRQBELNVTASKEPPDHVFPQIN 300
DB 241 TDGEKFGDPLGYEDVIVPEADREGVIRYVIGVDAFRSEKSRQBELNVTASKEPPDHVFPQIN 300
QY 301 NFEALKTIOQLRKIKPAIESTQTGSSSSPEHEMSQSGPSAAITNSGNPLLSITVGSVDWAG 360
DB 301 NFEALKTIOQLRKIKPAIESTQTGSSSSPEHEMSQSGPSAAITNSGNPLLSITVGSVDWAG 360
QY 361 GVPLYTSKEKSTFINMTRVDSMDNDAYLVGAALILNRVQSLVLGAPRYQHIGLVMFER 420
DB 361 GVPLYTSKEKSTFINMTRVDSMDNDAYLVGAALILNRVQSLVLGAPRYQHIGLVMFER 420
QY 421 ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDALVLI GAPHYETGREGQVSVCEL 480
DB 421 ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDALVLI GAPHYETGREGQVSVCEL 480
QY 481 PRGQARWQCDVAVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 481 PRGQARWQCDVAVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 600
QY 601 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLFVQKSTRDLREGQIQSVVT 660
DB 601 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLFVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKSTRQTVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
QY 721 SLVGTPLSAFQNLAPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFQNLAPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLVSVEKVSFELQNSQSRWRLACESASTEV 840
DB 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLVSVEKVSFELQNSQSRWRLACESASTEV 840
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTSRVMQHQYQVSNLQORSILPLVPLVPV 960
DB 901 QLELPVKYAVVMVYTSHGVSQTKYLNFTASENTSRVMQHQYQVSNLQORSILPLVPLVPV 960
QY 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1020
DB 961 RLNQTVIWDROPQVTFENSELSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1020
QY 1021 PFGIQEENATLKGNSLDFWYIKTSHNHLLIVSTAEILFNDVSFTLLPQAGAFVRSQTEF 1080
DB 1021 PFGIQEENATLKGNSLDFWYIKTSHNHLLIVSTAEILFNDVSFTLLPQAGAFVRSQTEF 1080
QY 1081 KVEPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFPKQYKQKDMWSEGGPPGABPQ 1137
DB 1081 KVEPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFPKQYKQKDMWSEGGPPGABPQ 1137

677	DB	LDLALDPELLSPRAATFOETKRNLSRVVUGLKAHCENFLLLPSCVDEUSVPIIILKLNK	738
721	QY	SLVCTPLSAFNGNRPVLAEDAQRLFTALPFEKXNCGNDNICDDLSITTFPSWLDCLVVG	780
737	DB	TLVGKPLAFNLPMLAALAAQRYFTASLPEKXNCGADHICQDNLGISPSPGLKSLVVG	796
781	QY	GPFRNVTVVVRNDGDSYQTVTFPPPLDLSRKUSTLQNSQRSWRLACESASSTEV	840
797	DB	SNLELNAEVMVMDGDSYGTITTFSPAGLSRYVAEGQKQGLSLHUTCDSAPVG--	854
841	QY	SCALKSTSCSINHPIFENSRSVFNINFDVDSKASLGNKLLKANVTSENNMPRTNKTEF	900
855	DB	SGQTWSTSCRINHILIFGGQAQITPLATFDVSPKAVLGDRILLITANVSSENNTPTSKTTF	914
901	QY	QLELPVKYAVVWVTSHGVSCTKYLNFAS-ENTSRVMQHOYQVSNLQORSLSLVLPLVP	959
915	DB	QLELPVKYAVVTVSSHEQPTKYLNPSESEKESHVAMHRYQVNNLQORDLPVSIWFVP	974
960	QY	VLNLQTVWDRPQVTPSENLSSTCHTERLPSSHSDFLAELRKPAVVMVCSLAVCQRIQCDI	1019
975	DB	VELNQAQVMDVYVSHQPNSLRCSSKSEKAPPASDFLAHIOKPNPLDCLSLAGCLRFCDV	1034
1020	QY	PFEGQIEFNATLKNGISFOWYIKTSHNILLIYSTAEILFENDSVFTLLPQOGAFVRSQTE	1079
1035	DB	PSFVSQZELDTFLKGNLSFGWVRQILQKVSVVSVASITFDTSVYSQLPGQEAFMRAQTT	1094
1080	QY	TKVEPFRVFNPLIIVSSVGGLLLLLALITLALYKLGFFPKQYKDKXSE	1128
1095	DB	TVLEKYKVHNPTLIIVSSIGGLLLALITAVLYKVGFPKQYKEMXEE	1143

RESULT 12
 US-09-350-259-4
 ; Sequence 4, Application US/09350259
 ; Patent No. US20020062008A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.
 ; APPLICANT: Van der Vieren, Monica
 ; TITLE OF INVENTION: NO. US20020062008A1el Human 2
 ; FILE REFERENCE: 27866735004
 ; CURRENT APPLICATION NUMBER: US/09/350,259
 ; CURRENT FILING DATE: 1999-07-08
 ; EARLIER APPLICATION NUMBER: 09/193,043
 ; EARLIER FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: 08/173,497
 ; EARLIER FILING DATE: 1993-12-23
 ; EARLIER APPLICATION NUMBER: 08/286,889
 ; EARLIER FILING DATE: 1994-08-05
 ; EARLIER APPLICATION NUMBER: 08/362,652
 ; EARLIER FILING DATE: 1994-12-21
 ; EARLIER APPLICATION NUMBER: 08/943,363
 ; EARLIER FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1163
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-350-259-4

Query Match	58.6%	Score 3446;	DB 9;	Length 1163;
Best local Similarity	60.8%	Pred. No. 3e-307;		
Matches 686;	Conservative 147;	Mismatches 290;	Indels 6;	Gaps 4;

QY	1	FNLDTENAMTFQENARGFQSVVQLQGRVVVVGAPQETVAANQORGLSYQCDYSTGSCBPI	60
DB	20	FNLDTEELTAFRVDSAGFCDVSVQYANSVWVVGAPQKIIAANQIGGLYQCCGYSTGACBPI	79
QY	61	RLOQVPVEAVNMSLGLSLAATTSPPOLLACGGPTVHTCSENTYVKGCLCFPGSNLRQOQK	120
DB	80	GLQVPPPEAVNMSLGLSLAATTSPPOLLACGGPTVHRECCGNMYLTGLCFLLGPT--OLTOR	137
QY	121	FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEPFISTVMBQLKSKTLFLSLMOYSEF	180

138 LFVSRQECPRQEQDITVFLIDSGSLSSRNFAATMMNFRAVISQFQRPSTQFSLQFQSNKF 197
181 RIHFTFKFQNNPNRSLIKITQTLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
198 QTHFTFEFRSTNPLSLLASVHQLQGTYYTATAIQNVVHFLFASYGARRDAIKILIVI 257
241 TDGKFKGDLGVEDVPEADREGVIRYVIGDAPRSEKSEKSEKSEKSEKSEKSEKSEKSEK 300
258 TDGKFKGDLGVEDVPEADREGVIRYVIGDAPRSEKSEKSEKSEKSEKSEKSEKSEKSEK 317
301 NFEALKTIONLREKIPAIETGTTGSSSSFEHMSQEGFSAITNSGPELLTGVSYDWAG 360
318 DFDALKTIONLREKIPAIETGTTGSSSSFEHMSQEGFSAITNSGPELLTGVSYDWAG 377
361 GVFLYTSKEKSTFTINMTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
378 GAFLYPPNMSPTFFINNSQENVMDRDSYLGSTELALMKGVQSLVGLGAPRYOHIGLVAMFR 437
421 QNTGHWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLGAPHYEOTRGQVSVCP 480
438 QVSRQWRKAEVIGTQIGSYFGASICSVDVDSNGSTDLVLGAPHYEOTRGQVSVCP 497
481 PRGQARWOCDAVLYGEOQPMGRFGAALTVDVNGDKLTVDVAIGAPGEEENRGAVILF 540
498 PRGWR-RWMCDAVLYGEOQPMGRFGAALTVDVNGDKLTVDVAIGAPGEEENRGAVILF 556
541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGAGQHVLLRSQ 600
557 HGVLPSPISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGAGQHVLLRSQ 616
601 PVLRYKAIEMNPREVAENFECNDVVYKKEAGVRVCLHVQKSTRDLRREGQIQSVVT 660
617 PVLWVGVSQMOFIPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLSQSVT 676
720 YDLALDSGRPHSRVAFNETKSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 720
677 LDALAPGLSPRAIFQSTKRSLSRVVGLKACENFNILLPSCVEDSVPIILRNF 736
721 SLVGTPLSAGNLRPLVLAEDAQRLFTALEPFFKNCNGNDNICDDLSITPFSMLDCLVVG 780
737 TLVGKELLAFRLNRLPMLAALAQRYPTASLPFEKNCQADHICQDNLGISFPLGLKSLVG 796
781 GRPEFNVTVVNDGSDSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 840
797 SNLELNAEVNMDGSDSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 854
841 SGALKETSCSINHPIPEENSEVTFTFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900
855 SQGTASTSCSINHPIPEENSEVTFTFDVDSKASLGNKLLKANVTSENNPRTNKTEF 914
901 QLELPKYAVMYVTSKGVSTKILNTAS-ENTSRVMOHYOVSNLQORSLPISLVFLVP 959
915 QLELPKYAVMYVTSKGVSTKILNTAS-ENTSRVMOHYOVSNLQORSLPISLVFLVP 974
960 VRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKPVNVNCIAVCORIQCDI 1019
975 VELNQAVMDVSHVQHPSPKSCSEKTAAPASDFLAHOKNPLVDCSLACGLRCDV 1034
1020 PFFGIQEFNATLKNLSPDWYIKTSHNHLIVSTAEIILPNDVSFTLLPQCGAFVRSQTE 1079
1035 PPSVQSEELDFTLKNLSPDWYIKTSHNHLIVSTAEIILPNDVSFTLLPQCGAFVRSQTE 1094
1080 TKVEPEFVNPPLIVGSSVGGILLALITAAVLYKLGPFKRYKQVKNWSE 1128
1095 TVLEKRYKVNPIPLIVGSSVGGILLALITAAVLYKLGPFKRYKQVKNWSE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007728A1

; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2003007728A1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1163
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.6%; Score 3446; DB 10; Length 1163;

Best Local Similarity 60.8%; Pred. No. 3e-307;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 PNLDTENAMTFQENARVGFQSVVOLQGRVVGAPQEIIVANQSGSLYQDYSTGSCPEI 60
DB 20 PNLDTENAMTFQENARVGFQSVVOLQGRVVGAPQEIIVANQSGSLYQDYSTGSCPEI 79
QY 61 RLQVPEAVNMSLGLSLAATTPPQLLACGPTVHOTSCTENTYVKGCLFGLFSGNLRRQOPK 120
DB 80 GLQVPPAVNMSLGLSLAATTPPQLLACGPTVHOTSCTENTYVKGCLFGLFSGNLRRQOPK 137
QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRRAKEPFTVNEQKKSKTLPFLMOYSEEP 180
DB 138 LPVSRQECPRQEQDITVFLIDSGSLSSRNFAATMMNFRAVISQFQRPSTQFSLQFQSNKF 197
QY 181 RIHFTFKFQNNPNRSLIKITQTLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
DB 198 QTHFTFEFRSTNPLSLLASVHQLQGTYYTATAIQNVVHFLFASYGARRDAIKILIVI 257
QY 241 TDGKFKGDLGVEDVPEADREGVIRYVIGDAPRSEKSEKSEKSEKSEKSEKSEKSEKSEK 300
DB 258 TDGKFKGDLGVEDVPEADREGVIRYVIGDAPRSEKSEKSEKSEKSEKSEKSEKSEKSEK 317
QY 301 NFEALKTIONLREKIPAIETGTTGSSSSFEHMSQEGFSAITNSGPELLTGVSYDWAG 360
DB 318 DFDALKTIONLREKIPAIETGTTGSSSSFEHMSQEGFSAITNSGPELLTGVSYDWAG 377
QY 361 GVFLYTSKEKSTFTINMTRVDSMDNAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFFINNSQENVMDRDSYLGSTELALMKGVQSLVGLGAPRYOHIGLVAMFR 437
QY 421 QNTGHWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLGAPHYEOTRGQVSVCP 480
DB 438 QVSRQWRKAEVIGTQIGSYFGASICSVDVDSNGSTDLVLGAPHYEOTRGQVSVCP 497
QY 481 PRGQARWOCDAVLYGEOQPMGRFGAALTVDVNGDKLTVDVAIGAPGEEENRGAVILF 540
DB 498 PRGWR-RWMCDAVLYGEOQPMGRFGAALTVDVNGDKLTVDVAIGAPGEEENRGAVILF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGAGQHVLLRSQ 600
DB 557 HGVLPSPISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVLTGAGQHVLLRSQ 616
QY 601 PVLRYKAIEMNPREVAENFECNDVVYKKEAGVRVCLHVQKSTRDLRREGQIQSVVT 660
DB 617 PVLWVGVSQMOFIPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLSQSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 LDALAPGLSPRAIFQSTKRSLSRVVGLKACENFNILLPSCVEDSVPIILRNF 736

APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: NO US20030077278a1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO. 2
LENGTH: 1161
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.0%; Score 3411; DB 10; Length 1161;
Best Local Similarity 59.4%; Pred. No. 5e-304;
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPOEIVAAANORGSLYKODIDTSGCEPI 60
DB 17 FNLDVEEPTIFQDADGFGOSVVOLOGSRVVGAPLEVAANOTGRXYDCAAAATGMCQPI 76
QY 61 RLOVPEAVNMISGLISAATTSPPOLLACQPTVHOTCSENTYVKGCLFLFGSNLRQPOK 120
DB 77 PLHIREAVNMISGLISATTSPPOLLACQPTVHOTCSENTYVKGCLFLFGSNLRQPOK 135
QY 121 FPEALRGCPQEDSDIAFLVDGSGIIPHPFRARKEFISTVMEQKSKTFLSLMQYSEEP 180
DB 136 VPDATPCPEQEMDIIVFLIDGSGIDQNDPQNMKGFGVAVMGQFEGDTLFLALMQYSNLL 195
QY 181 RIHFTPEKQONNPSRLIPIIQLLQRTHTATGIRKVVRELENIINGARKNAFKILILI 240
DB 196 KIHFTFQPTSPSQSLVDPIVOLKGLTFATGILTVVTFQFHKNKARKSAKILIVI 255
QY 241 TCGEKFGDPLGYEDVIEADREGVIRYVIGVGDFAFRSEKSRQBLNTVASKPPRDHVFQIN 300
DB 256 TDGQKYNDPLEYSDVIPAQEKAGIIRVAIGVGHAFQPTARQBLNTISSAPPQDHVFKVD 315
QY 301 NFEALKTIONLREKIPIAIGETQGTGSSSEHEMSQEGPSAAITSNCEPLLSTVGSYDAG 360
DB 316 NFAALGSTQKQEKIYAVEGTQSRASSSTQHEMSQEGFSTALTMDGLFLGAVGVSFSWSG 375
QY 361 GVELYTSKESKSTRINMTRVDSMDMDAYLVGAAAIILRNVRQSLVIGAPRYOHIGLVAMFR 420
DB 376 GAFLYPPNMGPTFLNMGQENVMDSDYLGSTELALAKGQNLVIGAPRYOHTGKAVIFT 435
QY 421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSDTLVIGAPHYTEOTRGQVSVCP 480
DB 436 QVSRQWRKKAQVETQIGSYFGASLCSVDVDSNGSDTLVIGAPHYTEOTRGQVSVCP 495
QY 481 PRGQARWQCDVLYGQGPQWGRFGAALTVLGVDNGDKLTVAGAPGEEDNRGAVYLF 540
DB 496 PRGQVQWQCDVLYGQGPQWGRFGAALTVLGVDNGDKLTVAGAPGEEDNRGAVYLF 555
QY 541 RGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGQGLTMDGLVLDLVGAGCHVLLRSQ 600
DB 556 HGASESGISPSHSQRIAGSKLSPRLQVFGQSLGQGLTMDGLVLDLVGAGCHVLLRSQ 615
QY 601 PYLRVKAIMEPNPREVARNFECDNDVVKGEAGEVRCVLCVOKSTRDLRREGOIQSVWT 660
DB 616 FVLKVGAVMFRFSPVEAVKAVRCWEEKPSALRAGDATVCLTIQKSSLDQL--GDIQSVR 673
QY 661 YDLALDGRPHSRVAFNFTKSTRQTLGLTQTCETLKLQLENFIEDPVPVILRLNF 720
DB 674 FDLALDPCGLATSRAIFNETKPTLTERKTLGLGTHCETLKLLEPCQVEDVVSPIILHNF 733

Search completed: June 7, 2004, 17:38:51
Job time : 42.8905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPO 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	1	US-08-173-497-3
2	5862	99.8	1153	1	US-08-286-889-3
3	5862	99.8	1153	1	US-08-485-618-3
4	5862	99.8	1153	1	US-08-362-652-3
5	5862	99.8	1153	2	US-08-605-672-3
6	5862	99.8	1153	2	US-08-482-293A-3
7	5862	99.8	1153	2	US-08-943-363-3
8	5862	99.8	1153	3	US-09-193-043-3
9	5862	99.8	1153	4	US-09-688-307A-3
10	5862	99.8	1153	4	US-09-350-259-3
11	5831.5	99.2	1152	2	US-08-476-062A-43
12	5831.5	99.2	1152	5	PCT-US96-01314-43
13	5831.5	99.2	1152	6	5424399-2
14	3469	59.0	1163	2	US-08-476-062A-44
15	3469	59.0	1163	5	PCT-US96-01314-44
16	3446	58.6	1163	1	US-08-173-497-4
17	3446	58.6	1163	1	US-08-286-889-4
18	3446	58.6	1163	1	US-08-485-618-4
19	3446	58.6	1163	1	US-08-362-652-4
20	3446	58.6	1163	2	US-08-605-672-4
21	3446	58.6	1163	2	US-08-482-293A-4
22	3446	58.6	1163	2	US-08-943-363-4
23	3446	58.6	1163	3	US-09-193-043-4
24	3446	58.6	1163	4	US-09-688-307A-4
25	3446	58.6	1163	4	US-09-350-259-4
26	3411	58.0	1161	1	US-08-173-497-2
27	3411	58.0	1161	1	US-08-286-889-2

28	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3411	58.0	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3411	58.0	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3411	58.0	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3411	58.0	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3411	58.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3411	58.0	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3395.5	57.8	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3395.5	57.8	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3395.5	57.8	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3395.5	57.8	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3395.5	57.8	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3395.5	57.8	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3395.5	57.8	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3232.5	55.0	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3232.5	55.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3232.5	55.0	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVAAQNGRGLYQCDYTGSGCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQEIIVAAQNGRGLYQCDYTGSGCEPI 76

QY 61 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRLOQPOK 120
DB 77 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRLOQPOK 136
QY 121 PPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKSKXTLFSLMQYSEEF 180
DB 137 PPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKSKXTLFSLMQYSEEF 196
QY 181 RHFTTFKPFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 240
DB 197 RHFTTFKPFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILRNRVOSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILRNRVOSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQVSVCP 480
DB 437 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQVSVCP 496
QY 481 PRGQARWOCDAVLGEOGQPRGCAALTVLGDVNGDKLTDVAIGAPESDNGAVYLF 540
DB 497 PRGQARWOCDAVLGEOGQPRGCAALTVLGDVNGDKLTDVAIGAPESDNGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVAQGHVLLRSQ 616
QY 601 PVLKVAIMEFPREVARVPCNDQVVKGEAGEVRVCLHVOKSTRDLRBEQIQTQVVT 660
DB 617 PVLKVAIMEFPREVARVPCNDQVVKGEAGEVRVCLHVOKSTRDLRBEQIQTQVVT 676
QY 661 YDLALDSGRPHSAFVNETKSTRQTVGLTQTCETILKQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSAFVNETKSTRQTVGLTQTCETILKQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPPFKNGCNDNICDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPPFKNGCNDNICDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTTVNDGSDSVRTQVTFPPDLDSYRKVSTLONORSQSWELACESASSTEV 840
DB 797 GPREFNVTTVNDGSDSVRTQVTFPPDLDSYRKVSTLONORSQSWELACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
QY 901 QLELPVKYAVMVTSHGVSATYKLAFTASENTSRVWQHGYQVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVMVTSHGVSATYKLAFTASENTSRVWQHGYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVWNCISIAVCORIQCIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVWNCISIAVCORIQCIP 1036
QY 1021 PFGIOBEFNATLKNLSFPMYIKTSINHLLIYSTABILFNDVSFTLLPQCGAFVRSOTET 1080
DB 1037 PFGIOBEFNATLKNLSFPMYIKTSINHLLIYSTABILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 XVEPFPVNPPLPLIVGSSVGGILLALITAAALYKLGFPKRYQKDWMSSEGGPFGABPQ 1137
DB 1097 XVEPFPVNPPLPLIVGSSVGGILLALITAAALYKLGFPKRYQKDWMSSEGGPFGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr, Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRLOQPOK 120
DB 77 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRLOQPOK 136
QY 121 PPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKSKXTLFSLMQYSEEF 180
DB 137 PPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKSKXTLFSLMQYSEEF 196
QY 181 RHFTTFKPFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 240
DB 197 RHFTTFKPFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIECTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

361 GVFLYTSKEKSTFNMTRVSDMDNDAYLGAAALILRNVRQSLVGLGAPRYQHIGLVAMFR 420
Db |
377 GVFLYTSKEKSTFNMTRVSDMDNDAYLGAAALILRNVRQSLVGLGAPRYQHIGLVAMFR 436
Qy |
421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSCPL 480
Db |
437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSCPL 496
Qy |
481 PRGQARWQCDVAVLYGEGQCPWGRFGAALTVDLVGVDNGDKLTDVAIGAPGEDNRGAVLYF 540
Db |
497 PRGQARWQCDVAVLYGEGQCPWGRFGAALTVDLVGVDNGDKLTDVAIGAPGEDNRGAVLYF 556
Qy |
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db |
557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy |
601 PVLRVKATMEFNPREVARVFNPCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
Db |
617 PVLRVKATMEFNPREVARVFNPCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 676
Qy |
661 YDLALDSGRPHSRAVFNETKSTRQTVGLGTOTCETLKLQLENCEIDPVSPVLRINF 720
Db |
677 YDLALDSGRPHSRAVFNETKSTRQTVGLGTOTCETLKLQLENCEIDPVSPVLRINF 736
Qy |
721 SLVGTPLSAFGNLRVLAEDAQRLEFALFPPEKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db |
737 SLVGTPLSAFGNLRVLAEDAQRLEFALFPPEKNGCNDNICODDLSITFSFMSLDCLVVG 796
Qy |
781 GPREFNVTTVVNDGEDSVRTQVTFPPFLDLSYRVKSTLQNRQSRVRLACESASSTEV 840
Db |
797 GPREFNVTTVVNDGEDSVRTQVTFPPFLDLSYRVKSTLQNRQSRVRLACESASSTEV 856
Qy |
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTENNMPRNKTEF 900
Db |
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTENNMPRNKTEF 916
Qy |
901 QLELPVKYAVVMVTSHGVSSTYKTLAFTASNTSRVMQHOYQVSNLQSRSLPISLVELVPV 960
Db |
917 QLELPVKYAVVMVTSHGVSSTYKTLAFTASNTSRVMQHOYQVSNLQSRSLPISLVELVPV 976
Qy |
961 RLNTQVIWDRPQVTFSENLSSTCHTKERLPUSDHFLAELRKAPVNCISIAVCQRIQCIP 1020
Db |
977 RLNTQVIWDRPQVTFSENLSSTCHTKERLPUSDHFLAELRKAPVNCISIAVCQRIQCIP 1036
Qy |
1021 PFCIQEEFNATLKNLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTF 1080
Db |
1037 PFCIQEEFNATLKNLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTF 1096
Qy |
1081 KVEPPEVPNPPLPIVSGSSVGGLLILALITAAALYKLGFPKRYQKDMWSEGGPPGABPQ 1137
Db |
1097 KVEPPEVPNPPLPIVSGSSVGGLLILALITAAALYKLGFPKRYQKDMWSEGGPPGABPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3

Query Match 99.84; Score 5862; DB 1; Length 1153;

Best Local Similarity 99.34; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy	1	PNLDENAMTPOENARGFGQSVVOLQGRVVVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI	60
Db	17	PNLDENAMTPOENARGFGQSVVOLQGRVVVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI	76
Qy	61	RLOVPVEAVNNSLGLSAAATSPPOLLAGCTVHTCTSENTYVKGCLFLFGSNLRQPOK	120
Db	77	RLOVPVEAVNNSLGLSAAATSPPOLLAGCTVHTCTSENTYVKGCLFLFGSNLRQPOK	136
Qy	121	FPEALRGCPQSDSIAPLVDGSGSIIIPHDFFRAKEFISTVMEQKKSKTLFSLMOYSEEP	180
Db	137	FPEALRGCPQSDSIAPLVDGSGSIIIPHDFFRAKEFISTVMEQKKSKTLFSLMOYSEEP	196
Qy	181	RIHFTPKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARNAFKLILII	240
Db	197	RIHFTPKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARNAFKLILVI	256
Qy	241	TGDKFKGDPGLGYEDVIPADREGVIRYVIGVGDAFRSEKSOBLNTVASKPPRDHVFQIN	300
Db	257	TGDKFKGDPGLGYEDVIPADREGVIRYVIGVGDAFRSEKSOBLNTIASKPPRDHVFQIN	316
Qy	301	NFEALKTIQNLREKIFAIETGTGSSSSPEHMSQBGFSAAITNSGILLSTVGSYDNAG	360
Db	317	NFEALKTIQNLREKIFAIETGTGSSSSPEHMSQBGFSAAITNSGILLSTVGSYDNAG	376
Qy	361	GVFLYTSKEKSTFNMTRVSDMDNDAYLGAAALILRNVRQSLVGLGAPRYQHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFNMTRVSDMDNDAYLGAAALILRNVRQSLVGLGAPRYQHIGLVAMFR	436
Qy	421	QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSCPL	480
Db	437	QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSCPL	496
Qy	481	PRGQARWQCDVAVLYGEGQCPWGRFGAALTVDLVGVDNGDKLTDVAIGAPGEDNRGAVLYF	540
Db	497	PRGQARWQCDVAVLYGEGQCPWGRFGAALTVDLVGVDNGDKLTDVAIGAPGEDNRGAVLYF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ	600

[illegible]

RESULT 4

US-08-1362-652-3
Sequence 3, Application US/0832652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

[illegible]

RESULT 5

```

1  US-08-605-672-3
2  Sequence 3, Application US/08605672
3  Patent No. 5817515
4  GENERAL INFORMATION:
5  APPLICANT: Gallatin, W. Michael
6  APPLICANT: Van der Vliet, Monica
7  TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
8  NUMBER OF SEQUENCES: 103
9  CORRESPONDENCE ADDRESS:
10 ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
11 STREET: 233 South Wacker Drive, 6300 Sear Tower
12 CITY: Chicago
13 STATE: Illinois
14 COUNTRY: United States
15 ZIP: 60608-6402
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release #1.0, Version #1.25
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/605,672
23 FILING DATE:
24 CLASSIFICATION: 530
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 08/173,497
27 FILING DATE: 23-DEC-1993
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 08/286,889
30 FILING DATE: 5-AUG-1994
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 08/362,652
33 FILING DATE: 21-DEC-1994
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Williams Jr., Joseph A.
36 REGISTRATION NUMBER: 38,659
37 REFERENCE/DOCKET NUMBER: 27866/32684
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: 312-474-6300
40 TELEFAX: 312-474-0448
41 TELEX: 25-3856
42 INFORMATION FOR SEQ ID NO: 3:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 1153 amino acids
45 TYPE: amino acid
46 STRANDEDNESS: single
47 TOPOLOGY: linear
48 MOLECULE TYPE: protein
49 US-08-605-672-3

```

US-08-605-672-3

Query Match	99.8%	Score 5862;	DB 2;	Length 1153;		
Best Local Similarity	99.3%;	Pred. No. 0;				
Matches 1129;	Conservative	7;	Mismatches 1;	Indels 0; Gaps 0;		
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSVVVVGAPQEI	VAANORGSLYQCDYSTGSC	EPI 60		
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSVVVVGAPQEI	VAANORGSLYQCDYSTGSC	EPI 76		
QY	61	RLQYPVAVNMSLGLSLAATTSPQALLACGPTVHOT	CSENTYVKGCLPFLFGSNLRQ	QOK 120		
DB	77	RLQYPVAVNMSLGLSLAATTSPQALLACGPTVHOT	CSENTYVKGCLPFLFGSNLRQ	QOK 136		
QY	121	PPREALRQCPQDSDIAPLDVGGSSII	PHDPRRAKEFI	STYMBOLKKSITPLSLMOYS	EEP 180	
DB	137	PPREALRQCPQDSDIAPLDVGGSSII	PHDPRRAKEFI	STYMBOLKKSITPLSLMOYS	EEP 196	
QY	181	RIHFTFKFQNNPNRSLKPIQTQLGRTH	TATGIRKVVRELFNITGAR	KNAPKILILI 240		
DB	197	RIHFTFKFQNNPNRSLKPIQTQLGRTH	TATGIRKVVRELFNITGAR	KNAPKILIVI 256		
QY	241	TQGEKKKDPGLGYEDVTP	PEADRGVTRYVIGCDAP	ERSKSRQELNTVAS	KPPRPHVPOIN 300	
DB	257	TQGEKKKDPGLGYEDVTP	PEADRGVTRYVIGCDAP	ERSKSRQELNTVAS	KPPRPHVPOIN 316	
QY	301	NFEALKTIONLREKIPAI	EGTQTSSSSFEHMSORGF	SAAITNSG	PLLTSTVGSYD	WAG 360
DB	317	NFEALKTIONLREKIPAI	EGTQTSSSSFEHMSORGF	SAAITNSG	PLLTSTVGSYD	WAG 376
QY	361	GVFLYTSKESKSTFINNTR	VDSOMDAYLG	YAAAIILNRVQSLV	LGAPRYOHIGL	VAMFR 420
DB	377	GVFLYTSKESKSTFINNTR	VDSOMDAYLG	YAAAIILNRVQSLV	LGAPRYOHIGL	VAMFR 436
QY	421	QNTQWHSNANVKGTOIGAY	FGASLCSVDVDSNG	STDVLITG	APHYBQTRGGQV	SVCP 480
DB	437	QNTQWHSNANVKGTOIGAY	FGASLCSVDVDSNG	STDVLITG	APHYBQTRGGQV	SVCP 496
QY	481	PRGQARWQCDAYLYG	EQGQPMGRFGAALT	VLGDVNGDKLTDVAI	GAPGEEDNRGA	VTYLF 540
DB	497	PRGQARWQCDAYLYG	EQGQPMGRFGAALT	VLGDVNGDKLTDVAI	GAPGEEDNRGA	VTYLF 556
QY	541	HGTSSGSI	SPSHSORLAGSKL	SPRLQYFQOSLS	GGODLTMDGLVDLT	TVGAQGHVLLARSQ 600
DB	557	HGTSSGSI	SPSHSORLAGSKL	SPRLQYFQOSLS	GGODLTMDGLVDLT	TVGAQGHVLLARSQ 616
QY	601	PVLKVAIMFNP	PREVARNVFECNDQV	KGKEAGBVRVCLHVQ	KS	TRDLRABGOIQSVVT 660
DB	617	PVLKVAIMFNP	PREVARNVFECNDQV	KGKEAGBVRVCLHVQ	KS	TRDLRABGOIQSVVT 676
QY	661	YDLALDSGRPHSR	AVNETKNSTRQTVL	GLTQCTETKLQ	LQPNCI	EDPVPSTVLRNLF 720
DB	677	YDLALDSGRPHSR	AVNETKNSTRQTVL	GLTQCTETKLQ	LQPNCI	EDPVPSTVLRNLF 736
QY	721	SLVGTPLSARGNLPVLA	EDAKELFTALPP	EPKNGCNDNICOD	DISIT	TFSPMSLDCLAVG 780
DB	737	SLVGTPLSARGNLPVLA	EDAKELFTALPP	EPKNGCNDNICOD	DISIT	TFSPMSLDCLAVG 796
QY	781	GPRBFNVTVTRND	GSDSYRTQVTFP	FDLDSYRKVSTLQ	NRQSRGNLR	CESASSTEV 840
DB	797	GPRBFNVTVTRND	GSDSYRTQVTFP	FDLDSYRKVSTLQ	NRQSRGNLR	CESASSTEV 856
QY	841	SGALKSTSCSINHPI	IPENSEVTFNIT	FDVDSKASLGNKLL	KANVTSENNAP	TKNTKTF 900
DB	857	SGALKSTSCSINHPI	IPENSEVTFNIT	FDVDSKASLGNKLL	KANVTSENNAP	TKNTKTF 916
QY	901	QLEIPVKYAVVMV	THSGHVSXTKL	APTASENTSRVMQ	HOYVSNL	GORSPLISLTVLPV 960
DB	917	QLEIPVKYAVVMV	THSGHVSXTKL	APTASENTSRVMQ	HOYVSNL	GORSPLISLTVLPV 976
QY	961	RLNQTIVWDR	POQVTFSEN	SSCHTKERLP	SHSDFLAELRKAP	VVNCIAVCQRIQCDIP 1020
DB	977	RLNQTIVWDR	POQVTFSEN	SSCHTKERLP	SHSDFLAELRKAP	VVNCIAVCQRIQCDIP 1036

Query March 99.8%: score 5862: DB 2: Length 1153:

Query match 53.8%; score 3882; DB 2; benign 1133;
Best Local Similarity 99.3%; pred. No. 0:

Best local similarity 99.5%, tied: no. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

1. The first part of the paper is devoted to the study of the asymptotic behavior of the solutions of the system (1) as $t \rightarrow \infty$. It is shown that the solutions of the system (1) are bounded and tend to zero as $t \rightarrow \infty$ if the matrix A is stable. The second part of the paper is devoted to the study of the asymptotic behavior of the solutions of the system (1) as $t \rightarrow \infty$ if the matrix A is not stable. It is shown that the solutions of the system (1) are unbounded and tend to infinity as $t \rightarrow \infty$ if the matrix A is not stable.

1 FNLDTENAMTFQENARGFGOSVWLOGSRVWVGAPQBIVAANORGSLYQCDYSTGSCEPI 60

17 FNLDTENAMTFQENARGFGOSVVOLOGSRVVVGAPOEIVAANQRGSLYQCDYSTGSCEPI 76

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGLCFLFGSNLRQQPQK 120

77 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCFLFGSNLRQQPOK 136

1. The first part of the document is a list of the names of the members of the committee, which is headed by the Chairman, Mr. J. H. M. J. van der Meulen. The list includes the names of the members of the committee, the names of the members of the sub-committee, and the names of the members of the working group.

121 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEOLKSKTLFSLMOYSEEF 180

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

137 FPEALRGCPQEDSDIAFLIDCGSGSIIPHDFFRRMKFEFVSTVMBOLKKSKTLFSLMOYSEEF 196

THE UNIVERSITY OF CHICAGO

181 RIHFTFKEFONNPNSRLIKPITQLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240

[illegible]

197 RIHETFKEFFONNPNRSLVKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256

[illegible]

241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPRPDHVFOIN 300

.....

257 TDGEKFGDPLGYBDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPDRDHVPQVN 316

1. The first step in the process of creating a new product is to identify a market need. This involves conducting market research to understand the preferences and behaviors of potential customers. Once a need is identified, the next step is to develop a concept that addresses this need. This concept should be unique, valuable, and feasible. The third step is to create a prototype, which is a preliminary model of the product. This allows the team to test the concept and make necessary adjustments. The fourth step is to conduct a feasibility study, which evaluates the technical, financial, and operational aspects of the product. Finally, the product is launched into the market, and the team monitors its performance and customer feedback to make further improvements.

301 NFEALKTIONQLREKIFAIEGTQTGSSSFEHMSQGFSAITSNGLPLSTVGSYDWAG 360

317 NREALKTIONQLREKI FAIEGTQTGSSSEHEMSOEGFSAATSNGLPLLS TVGSYDWAG 376

1. The first step in the process of creating a new product is to identify a market need. This involves conducting market research to understand the preferences and behaviors of potential customers. Once a need is identified, the next step is to develop a concept that addresses this need. This concept should be unique and offer a clear value proposition to the target market.

361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRRNVQSLVLGAPRYQHIGLVAMER 420

QY 1021 FPGIQEENATLKNLSFOWYIKTSNNHLLIVSTABILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1037 FPGIQEENATLKNLSFOWYIKTSNNHLLIVSTABILFNDVSFTLLPQCGAFVRSQTEF 1096
QY 1081 KVEPFEVNPPLVIGSSVGGLLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLVIGSSVGGLLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3
Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTPOENARGFGQSVVQLQGSRRVVVGAPQEIIVANQRGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTPOENARGFGQSVVQLQGSRRVVVGAPQEIIVANQRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEANMISGLSLAATTPSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQQPQK 120
Db 77 RLQVPVEANMISGLSLAATTPSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQQPQK 136
QY 121 FPEARLGCQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMBQLKSKTLFSLMOYSEEF 180

Db 137 FPEARLGCQEDSDIAFLVDGSGSIIPHDFFRMEKFYSTVNEQLKSKTLFSLMOYSEEF 196
QY 181 RIHTFKEFQNNPRLSIKPIITQLGRTHATGIRKVVRELFNITNGARNAPKILILI 240
Db 197 RIHTFKEFQNNPRLSIKPIITQLGRTHATGIRKVVRELFNITNGARNAPKILIVI 256
QY 241 TDGKFGDPLGYEDVPEADREGVIRVYIGVDGAFRSEKSHQELNLTASKPRPHVFOIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRVYIGVDGAFRSEKSHQELNLTASKPRPHVFOIN 316
QY 301 NFEALKTIONOLREKI PAIEGTQFGSSSPHEMSORGFSAITNSGFLLTGVSIDWAG 360
Db 317 NFEALKTIONOLREKI PAIEGTQFGSSSPHEMSORGFSAITNSGFLLTGVSIDWAG 376
QY 361 GVFLYTSKEKSTFINMTTRVDSMDNDAYLVYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTTRVDSMDNDAYLVYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYETQRTGGOVSVCPL 480
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYETQRTGGOVSVCPL 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNREGAVLYF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNREGAVLYF 556
QY 541 KTSSTSGISPSHSORIASGLSPLOYFGOSLGGQDLTMDCLVDLTVGAQHVLILRSQ 600
Db 557 HGTSGSGISPSHSORIASGLSPLOYFGOSLGGQDLTMDCLVDLTVGAQHVLILRSQ 616
QY 601 PVLAVKALMBENPREVARNVPECNDQVVKGEAGEVRLHVQVQSTDRDLREGOIQSVVT 660
Db 617 PVLAVKALMBENPREVARNVPECNDQVVKGEAGEVRLHVQVQSTDRDLREGOIQSVVT 676
QY 661 YDLALDSGRPHSRVAVFETKNSTRQOVGLTQCTETLKLQPNCEIDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVFETKNSTRQOVGLTQCTETLKLQPNCEIDPVSPVILRLNF 736
QY 721 SLVGTPLSAFGLNLPVLAQALFTALFPPEKNCNDNICODDLSITTFSPMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLPVLAQALFTALFPPEKNCNDNICODDLSITTFSPMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
QY 901 QLELPVKYAVYVTVVTSRGTSTKYLNFPTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVYVTVVTSRGTSTKYLNFPTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISVACQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVWNCISVACQRIQCDIP 1036
QY 1021 PFGIQEENATLKNLSFOWYIKTSNNHLLIVSTABILFNDVSFTLLPQCGAFVRSQTEF 1080
Db 1037 PFGIQEENATLKNLSFOWYIKTSNNHLLIVSTABILFNDVSFTLLPQCGAFVRSQTEF 1096
QY 1081 KVEPFEVNPPLVIGSSVGGLLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLVIGSSVGGLLALITAAALYKLGFFKQYKDMXSEGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
TYPE: amino acid
LENGTH: 1153 amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAQPIVAANQSGSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAQPIVAANQSGSLYQCDYSTGSCPEI	76
Qy	61	RLOVPVAVNMSLGLSUAATSPOLLACGPTVHQTSENTRYKGLCFLFGSNLRQOPQK	120
Db	77	RLOVPVAVNMSLGLSUAATSPOLLACGPTVHQTSENTRYKGLCFLFGSNLRQOPQK	136
Qy	121	FPBALGCPQSDIAFLVDGSGIIPHDSPRAKEFTVMEQLKSKTLFSLMOYSEEF	180
Db	137	FPBALGCPQSDIAFLVDGSGIIPHDSPRAKEFTVMEQLKSKTLFSLMOYSEEF	196
Qy	181	RIHPTFEFQNNPRLSIKEITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240
Db	197	RIHPTFEFQNNPRLSIKEITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI	256
Qy	241	TGKEFGDPLGYEDVIEADREGVIRVVGVDAPFRSEKSRQELNTVASKPRDHVQIN	300
Db	257	TGSKFGDPLGYEDVIEADREGVIRVVGVDAPFRSEKSRQELNTVASKPRDHVQVN	316
Qy	301	NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQSGFSAATSNGLPILSTVGSYDNAG	360
Db	317	NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQSGFSAATSNGLPILSTVGSYDNAG	376

Qy	361	GVPLTYSKEKSTPINMTRVDSMDNDAYLGYAAAILILANRVQSLVLAGPRYOHICLVAMER	420
Db	377	GVPLTYSKEKSTPINMTRVDSMDNDAYLGYAAAILILANRVQSLVLAGPRYOHICLVAMER	436
Qy	421	QNTGMWESNANVRGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQEGGVSVCP	480
Db	437	QNTGMWESNANVRGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQEGGVSVCP	496
Qy	481	PRGORAEWQCDVLYGBOGPNCRFGAALTVDVWGDGKLTDAIAGPREDNRGAYTLF	540
Db	497	PRGORAEWQCDVLYGBOGPNCRFGAALTVDVWGDGKLTDAIAGPREDNRGAYTLF	556
Qy	541	HGTSGSGISPSHQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVLDLTGGAQGHVLLLSRQ	600
Db	557	HGTSGSGISPSHQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVLDLTGGAQGHVLLLSRQ	616
Qy	601	PVLAVKALMEFNPREVARNVFECDQVWKEAGEVRVCLHVQKSTRDRLEBQIQSVVT	660
Db	617	PVLAVKALMEFNPREVARNVFECDQVWKEAGEVRVCLHVQKSTRDRLEBQIQSVVT	676
Qy	661	YDLALDSGRPHSRAVFNENKSTRRTQVGLGTQTCETLKLQLPNCIEDPVSPIVLRNLF	720
Db	677	YDLALDSGRPHSRAVFNENKSTRRTQVGLGTQTCETLKLQLPNCIEDPVSPIVLRNLF	736
Qy	721	SLVGTPLSAPGNLRPVLAEDAQRLFTALPPFKNCGNDNICQDDLSITTFPMGLDCLVVG	780
Db	737	SLVGTPLSAPGNLRPVLAEDAQRLFTALPPFKNCGNDNICQDDLSITTFPMGLDCLVVG	796
Qy	781	GPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNBSSQSRWLACASASTEV	840
Db	797	GPREFNVTVVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNBSSQSRWLACASASTEV	856
Qy	841	SGALKSTSCSINHPPIPPENSEVTFNITPDVDSKASIGNKLLKLLKANVTSENKMPRTNTEP	900
Db	857	SGALKSTSCSINHPPIPPENSEVTFNITPDVDSKASIGNKLLKLLKANVTSENKMPRTNTEP	916
Qy	901	QLELPVKYAVYVTVTSHGVSTKYLNFTASENTSRVWQHGVQVSNLQORSILPISLVFLVPV	960
Db	917	QLELPVKYAVYVTVTSHGVSTKYLNFTASENTSRVWQHGVQVSNLQORSILPISLVFLVPV	976
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCIAVCQRIQCDIP	1020
Db	977	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCIAVCQRIQCDIP	1036
Qy	1021	FFGIQEEFNATLKGNSLFDWYIKTSINHLITVSTABILNDSVFTLLPGQCAFVRSQTET	1080
Db	1037	FFGIQEEFNATLKGNSLFDWYIKTSINHLITVSTABILNDSVFTLLPGQCAFVRSQTET	1096
Qy	1081	KVEPFEVNPPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO	1137
Db	1097	KVEPFEVNPPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPO	1153

RESULT 8
US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03


```
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match      99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTPOENARGFGOSVVOLOGSRVVGAPQEIIVAAANORGLSYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTPOENARGFGOSVVOLOGSRVVGAPQEIIVAAANORGLSYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 136

QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRAKFPISTVMQKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDPRAKFPISTVMQKSKTFLSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVIPEADREGVIRVIVGVDGAFRSEKSRQELNLTIVASKPPDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEADREGVIRVIVGVDGAFRSEKSRQELNLTIVASKPPDRHVFQIN 316

QY 301 NFEALKTIONQUREKIFALEGTQCTSSSFHEHMQEGFSAITNSGPIILSTVGSYDNAG 360
DB 317 NFEALKTIONQUREKIFALEGTQCTSSSFHEHMQEGFSAITNSGPIILSTVGSYDNAG 376

QY 361 GVLPTSKEKSTFINNTRVDSMDNAYLGAYAAAILRNVRQSLVLGAPRYQHIGLVAMPR 420
DB 377 GVLPTSKEKSTFINNTRVDSMDNAYLGAYAAAILRNVRQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSDTDLVILGAPHYVETRGQVSVCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSDTDLVILGAPHYVETRGQVSVCP 496

QY 481 PRGQARWQCDVLYGEOQPPWGRFGAALTVLGVDNGDKLTDVAIGAPCEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQPPWGRFGAALTVLGVDNGDKLTDVAIGAPCEEDNRGAVYLF 556

QY 541 HGTSGSISPSHSQRIAGSKLSPRLOYFGOSLSGGODLTMDGLVLDITVGAQGHVILRSQ 600
DB 557 HGTSGSISPSHSQRIAGSKLSPRLOYFGOSLSGGODLTMDGLVLDITVGAQGHVILRSQ 616

QY 601 PVLVRKAIMFNPREVARNVFECDNVKGEAGEVRVCLHVOKSTRDRLRGQIQSVVT 660
DB 617 PVLVRKAIMFNPREVARNVFECDNVKGEAGEVRVCLHVOKSTRDRLRGQIQSVVT 676

QY 661 YDLALDSGRPHSAFNETNSTRTQVGLTQTCETLKLQLPNCIEDPUSPIVLRNLF 720
DB 677 YDLALDSGRPHSAFNETNSTRTQVGLTQTCETLKLQLPNCIEDPUSPIVLRNLF 736

QY 721 SLVGTPLSAFGLNLRPVLAADAQRLFTALPPFEKNCNDNI CODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAADAQRLFTALPPFEKNCNDNI CODDLSITFSFMSLCLVVG 796

QY 781 GPRBNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQONRSORSWRLACESASSTEV 840
DB 797 GPRBNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQONRSORSWRLACESASSTEV 856

QY 841 SGALKSTSCS INHPIIPENSEVTNITFDVDSKASIGNKLLKANVTSENNPRNKTKEF 900
DB 857 SGALKSTSCS INHPIIPENSEVTNITFDVDSKASIGNKLLKANVTSENNPRNKTKEF 916

QY 901 QLELPVKYAVYVWVTVSHGVSTKYLNFPTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
```

Db 317 NFEALKTIONQREKIFAIEGTQCTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTVSDMDNDAYLGAAAILRNVRQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTVSDMDNDAYLGAAAILRNVRQSLVILGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVILGAPRYOHIGLVAMFR 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVILGAPRYOHIGLVAMFR 496
Qy 481 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLVRKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFFPKNGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFFPKNGNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYKRVSTLQNRORSORWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYKRVSTLQNRORSORWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENNMPRTNKTFF 900

Db 317 NFEALKTIONQREKIFAIEGTQCTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTVSDMDNDAYLGAAAILRNVRQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTVSDMDNDAYLGAAAILRNVRQSLVILGAPRYOHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVILGAPRYOHIGLVAMFR 480
Db 437 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVILGAPRYOHIGLVAMFR 496
Qy 481 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLVRKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCEIDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFFPKNGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFFPKNGNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYKRVSTLQNRORSORWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYKRVSTLQNRORSORWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTENITFQVDSKASLGNKLLKANVTSENNMPRTNKTFF 900

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 916
Qy 901 QLELPVKAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISIVFLVPV 960
Db 917 QLELPVKAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISIVFLVPV 976
Qy 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCDIP 1020
Db 977 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCDIP 1036
Qy 1021 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1080
Db 1037 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1096
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAALYKLGFFKQYKQVMSGGPPGAEPO 1137
Db 1097 KVEPEVNPPLIIVGSSVGGLLLLALITAALYKLGFFKQYKQVMSGGPPGAEPO 1153

RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476, 062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLOQSRVVVGAPOEIVAAQORGSILYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOQSRVVVGAPOEIVAAQORGSILYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLAACGPTVHQTCSENTYVKGCLFPLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLAACGPTVHQTCSENTYVKGCLFPLFGSNLRQOPQK 136
Qy 121 FPEALRCPOBDSIDIAFLVDGSGSIIIPDPRRAKEPISTVMBOLKKSKTLFSLMQYSSEF 180
Db 137 FPEALRCPOBDSIDIAFLVDGSGSIIIPDPRRAKEPISTVMBOLKKSKTLFSLMQYSSEF 196
Qy 181 RIHFTTFKFNPNPRSLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RIHFTTFKFNPNPRSLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 256
Qy 241 TDGKFGEDPLGYEDVITPEADREBQVIRVYVIGVDAFRSEKSRQELNLTVAASKPPRPHVQIN 300
Db 257 TDGKFGEDPLGYEDVITPEADREBQVIRVYVIGVDAFRSEKSRQELNLTVAASKPPRPHVQIN 316
Qy 301 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLISTVGSYDWAG 376
Qy 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAALILRNVSQSLVIGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAALILRNVSQSLVIGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKCTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYVYEQTRGGQVSVCP 480
Db 437 QNTGWESNANVKCTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYVYEQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGBQGPWGFAGALTVLGDVNGDKLTDVAIGAPGEDNAGAVLF 540
Db 497 PRG-RARWQCDVLYGBQGPWGFAGALTVLGDVNGDKLTDVAIGAPGEDNAGAVLF 555
Qy 541 HGTSGSISPSHSORIKAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSORIKAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLRVKALMEFNPREVARNFECDQVWVGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKALMEFNPREVARNFECDQVWVGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTOVLGTQTCETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNRPVLAEDAQRLFTALPPEKNGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPRENVTWVRNDEGDSVTOVTFEPLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
Db 796 GPRENVTWVRNDEGDSVTOVTFEPLDLSYRKVSTLQNRORSORSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 900
Db 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEP 915
Qy 901 QLELPVKAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISIVFLVPV 960
Db 916 QLELPVKAVYVMTSHGVSTKYLNTASENTSRVMQHQYQVSNLQSRSLPISIVFLVPV 975
Qy 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCDIP 1020
Db 976 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCDIP 1035
Qy 1021 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1080
Db 1036 PFGIOBEFNATLKGMLSPDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQCGAFVRSQTET 1095
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAALYKLGFFKQYKQVMSGGPPGAEPO 1137

Db 1096 KVEPEVENPFLIVGSSVGGILLALITAAALYKLGFFKQYKQVXMMSEGGPGABPQ 1152

RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnabou
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.2%; Score 5831.5; DB 5; Length 1152;
Best Local Similarity 99.0%; Pred No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBFI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBFI 76

QY 61 RLOVPVEAVNMVSLGLSLAATTSPPQLACGPTVHQTCSENVYKGLCLFLFGSNLRQOPQK 120
Db 77 RLOVPVEAVNMVSLGLSLAATTSPPQLACGPTVHQTCSENVYKGLCLFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEQLKSKTLFSLMQYSEBP 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFISTVMEQLKSKTLFSLMQYSEBP 196

QY 181 RIHFTKEFQNNPNSRLIPIITOLLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
Db 197 RIHFTKEFQNNPNSRLIPIITOLLGRTHATGIRKVVRELENTINGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFPQIN 300
Db 257 TDGEKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFPQIN 316

QY 301 NFEALKTIQNLREKIPFAIGTQTGSSSSPHEHMSQEGFSAATISNGPLISTVGSYDAG 360
Db 317 NFEALKTIQNLREKIPFAIGTQTGSSSSPHEHMSQEGFSAATISNGPLISTVGSYDAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVGLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVGLGAPRYQHIGLVAMPR 436

QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIYQTRGGQSVQCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIYQTRGGQSVQCP 496

QY 481 PRGORARWCCDAVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGBDNRCGAVYLF 540
Db 497 PRG-RARWCCDAVLYGEOQWGFAGALTIVLGDVNGDKLTDVAIGAPGBDNRCGAVYLF 555

QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQSHVLLLRQ 600
Db 556 HGTSGSGISPSHSORIASKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQSHVLLLRQ 615

QY 601 PVLRVKALIMBNPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 616 PVLRVKALIMBNPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 675

QY 661 YDLALDSGRPHSRVAFNETKSTRROTQVLGLTOTCETLKLQLPNCIBDPVSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRROTQVLGLTOTCETLKLQLPNCIBDPVSPVILRLNF 735

QY 721 SLVGTPLSAFGNLRPVLAEQAQRLFTALPPEKNGCNDNICODDLSITPFSKSLDCLVVG 780
Db 736 SLVGTPLSAFGNLRPVLAEQAQRLFTALPPEKNGCNDNICODDLSITPFSKSLDCLVVG 795

QY 781 GPREFNVTVVRNDGEDSDYRTQVTFPPFLDLSYRKVSTLQORQSRWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSDYRTQVTFPPFLDLSYRKVSTLQORQSRWRLACESASSTEV 855

QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTENNMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTENNMPRTNKTEF 915

QY 901 QLELPVKYAVVMVVTSEGVSTKYLAFNTASENTSRVWQHQQYQVSNLQORSLPISLVFLVPV 960
Db 916 QLELPVKYAVVMVVTSEGVSTKYLAFNTASENTSRVWQHQQYQVSNLQORSLPISLVFLVPV 975

QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLABELKAPVWVNCISIAVCORIQCDIP 1020
Db 976 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLABELKAPVWVNCISIAVCORIQCDIP 1035

QY 1021 PFGIOEEFNATLKGNLSPDWIKTSHNHLIIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1080
Db 1036 PFGIOEEFNATLKGNLSPDWIKTSHNHLIIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1095

QY 1081 KVEPEVENPFLIVGSSVGGILLALITAAALYKLGFFKQYKQVXMMSEGGPGABPQ 1137
Db 1096 KVEPEVENPFLIVGSSVGGILLALITAAALYKLGFFKQYKQVXMMSEGGPGABPQ 1152

RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2:
; LENGTH: 1152
5424399-2

Query Match 99.2%; Score 5831.5; DB 6; Length 1152;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPFOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPFOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 136

QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKSKTLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNRSLIKPTIQLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPTIQLLGRTHATGIRKVVRELFTNGARKNAFKILIVI 256

QY 241 TDGKFGDPLGVEDVPEADREGVIRYVIGVDDAFRSEKSRQELMTVASKPRDHVFQIN 300
DB 257 TDGKFGDPLGVEDVPEADREGVIRYVIGVDDAFRSEKSRQELMTVASKPRDHVFQIN 316

QY 301 NFEALKTIONLREKIFATEGTOTSSSFHEHMSQEGFSAATISNGPLLSVTGSDYDAG 360
DB 317 NFEALKTIONLREKIFATEGTOTSSSFHEHMSQEGFSAATISNGPLLSVTGSDYDAG 376

QY 361 GVFLYTSKEKSPFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSPFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLICAPHYVQTRGGQSVQVCL 480
DB 437 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLICAPHYVQTRGGQSVQVCL 496

QY 481 PRGORARWOCDAVLGYEQOPGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWOCDAVLGYEQOPGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVYLF 555

QY 541 HGTSGSISPSHSORJAGSKLSPRLQYFGOSLSCGODLTMGDLVLTGVAQGHVLLRSQ 600
DB 556 HGTSGSISPSHSORJAGSKLSPRLQYFGOSLSCGODLTMGDLVLTGVAQGHVLLRSQ 615

QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLRHVQKSTRDLRSGQIQSVVT 660
DB 616 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLRHVQKSTRDLRSGQIQSVVT 675

QY 661 YDLALDSGRPHSRVNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720
DB 676 YDLALDSGRPHSRVNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 735

QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALPFFPKNCNDNIQDDLSITPSFMSLDCLVVG 780
DB 736 SLVGTPLSAFNLRPVLAEDAQLFTALPFFPKNCNDNIQDDLSITPSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGSDSYRTQTFPPFLDLSYRKVSTLQNRQSRQSWFLACESASSTEV 840
DB 796 GPRESNVTVVRNDGSDSYRTQTFPPFLDLSYRKVSTLQNRQSRQSWFLACESASSTEV 855

QY 841 SGALKSTCSINHPIIPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNVPRNTKTEF 900
DB 856 SGALKSTCSINHPIIPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNVPRNTKTEF 915

QY 901 QLELPVKIAYVMVTSHGVSSTKYNFTASENTSRVMQHQYQVSNLQORSPLISLFLVFPV 960
DB 916 QLELPVKIAYVMVTSHGVSSTKYNFTASENTSRVMQHQYQVSNLQORSPLISLFLVFPV 975

QY 961 RLNQTIVDRPQVTFSENLSSTCHTERLPSSDPLAELRKAPVNCSTAVCORQCDDIP 1020
DB 976 RLNQTIVDRPQVTFSENLSSTCHTERLPSSDPLAELRKAPVNCSTAVCORQCDDIP 1035

QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTEF 1080

DB 1036 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTEF 1095

QY 1081 KVPFPEVNPFLPLIVGSSVGGLLILALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1137
DB 1096 KVPFPEVNPFLPLIVGSSVGGLLILALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1152

RESULT 14
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Atrout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-062A-44

Query Match 59.0%; Score 3469; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.9e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 20 FNLDTENAMTFOENARGFGQSVVQLQGGSRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 79

QY 61 RLQVPVEAVNMSLGLSLAATTSPFOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 120
DB 80 RLQVPVEAVNMSLGLSLAATTSPFOLLACGPTVHTCSENTYVKGCLFPLGSLNRQQPOK 137

QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKSKTLFSLMOYSEEF 180
DB 138 LPVRSQECPRQDILVFLIDGSGSISSNFATMNFRAVISQRPSTQSLQFSNKP 197

QY 181 RIHFTFKFQNNPNRSLIKPTIQLLGRTHATGIRKVVRELFTNGARKNAFKILILI 240

138 QTHFTFEFRFRTSNPLSLASVHQLQGGFTYTATATQNVVHLFASYGARRDATKILIVI 257
241 TDGKFGDPLGVEDVPEADREGVTVYVIGVGDAPRSEKSEQLNTVASKPRDHVQIN 300
258 TDGKFGDPLGVEDVPEADREGVTVYVIGVGDAPRSEKSEQLNTVASKPRDHVQIN 317
301 NPEALKTQNLQREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGELLTVGSYDWAG 360
318 DFDALKDQNLQREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGELLTVGSYDWAG 377
361 GVFLYTSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHTGKAVFT 420
378 GAFLYPPNMSPTFINMSQENVMDSDYLGYSYTELALWGVQSLVGLGAPRYOHTGKAVFT 437
421 QNTGWMESNANVKTQIGAYFGASILCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 480
438 QVSRQWRMKAETVGTQIGSYFGASILCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 497
481 PRGQARWCCDVLVYGEQGWGFRGAALTVLGDVNGDKLTDVAIGAPGBENRGAVYLF 540
498 PRGMR-RWMCDAVLYGEQGWGFRGAALTVLGDVNGDKLTDVAIGAPGBENRGAVYLF 556
541 HTSGSGISPSHSORIASKLSPLQYFGQSLSGGDLTMDGLVLTGAGQHVLLRSQ 600
557 HGVLPSPISPSHSORIASKLSPLQYFGQSLSGGDLTMDGLVLTGAGQHVLLRSQ 616
601 PVLRYKALMEFNPVARNVFECDQVVKGEAGVRVCLHVOKSVOKSTRLRREGIQSVVT 660
617 PVLWGVSNQFTPAISAPAFCEQVWSEQTLVQSNICLYIDKSKNLLGSRDIQSSVT 676
661 YDLALDSGPHSRAVFNKSTRTQVLTQTCETLKLQLPNCIEDPVSPVLRUNF 720
677 LBLALDPGLSPRAFTQTKNSLSRVRVIGVKAHCENFNLPLSCVEDSVPTITLRNF 736
721 SLVGTPLGAFGNLRLPVLADQRLFTALPPFKNGCNDNICODULSITFSMSLDCLVVG 780
737 TLVGPPLAFNLRLPVLADQRLFTALPPFKNGCNDNICODULSITFSMSLDCLVVG 796
781 GPRBNVTVVNDGDSVTRQVTFEPLDLSYKVSITLONQORSORSRLACESASSTEV 840
797 SNLENAEVMWVNDGDSVTRQVTFEPLDLSYKVSITLONQORSORSRLACESASSTEV 854
841 SGALKSTCSINHEPPEPENSEVTFNITFDVDSKASLGKXKLLKANVTSENMPRTNKTEP 900
855 SGTWSTSCINLIFRGAQITFLATFDVSPKAVLGBRLITANVSSENTPRTSKTIF 914
901 QLELPVKYAVVTVSHGVSTKYVNTAS-ENTSRVMOHQYOVUSMLGQSLDISLVFLVP 959
915 QLELPVKYAVVTVSHGVSTKYVNTAS-ENTSRVMOHQYOVUSMLGQSLDISLVFLVP 974
960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPVVNCSIAVCORIQCDI 1019
975 VELNQAVMMDVEVSHQVPSLRCSSEKIAPPASDFLAHQNPVLDLCSIACLPRCDV 1034
1020 PFGIOEENFATLKNLSFDWIKTSHNLLIIVSTAEILPNDSPVFTLLPGQAPVRSQTE 1079
1035 PFSVQGEELDTLKNLSFGWRIQKVVSVVAEITFDTSVYSQLPQGEAFNRQTT 1094
1080 TKVPEPEVNPPLIVGSSVGGHLLIILALITVLYKLGPEKRYQDMWSE 1128
1095 TVLEKYKHVNPPLIVGSSVGGHLLIILALITVLYKLGPEKRYQDMWSE 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44

Query March 59.08; Score 3469; DB 5; Length 1163;
Best Local Similarity 61.04; Pred. No. 5.9e-284; Indels 6; Gaps 4;
Matches 689; Conservative 141; Mismatches 293;

QY 1 FNLDENAMTPOENARGFGQSVVQLQSGRVVVGAPQRIIVANQSGSLYQCVSTGSCBPI 60
DB 20 FNLDTEELTAFRVDSAGFSDVQVYANSVVGAPQRIIVANQSGSLYQCVSTGSCBPI 79
QY 61 RLQVPEAVNMSLGLSLAATSPQLLACGPTVHOTSENVVYVGLCLPFLGSLNRQQPK 120
DB 80 GLQVPEAVNMSLGLSLAATSPQLLACGPTVHOTSENVVYVGLCLPFLGSLNRQQPK 137
QY 121 FPEALGCPQSDSDIAFLVDGSGSIIIPHDPRRAKEFISTVMEQLKSKTLFSLMOYSSEF 180
DB 138 LPVSRQECPRQEQIIVFLDQSGSISRNFAFGVAVISQFQRPSTQFSLAQFSNK 197
QY 181 RIHTTFKFPQNNPRLIKPITQLLGHRTATGIRKVVRELFTNAGRNAPKILILI 240
DB 198 QTHFTFEFRFRTSNPLSLASVHQLQGGFTYTATATQNVVHLFASYGARRDATKILIVI 257
QY 241 TDGKFGDPLGVEDVPEADREGVTVYVIGVGDAPRSEKSEQLNTVASKPRDHVQIN 300
DB 258 TDGKFGDPLGVEDVPEADREGVTVYVIGVGDAPRSEKSEQLNTVASKPRDHVQIN 317
QY 301 NPEALKTQNLQREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGELLTVGSYDWAG 360
DB 318 DFDALKDQNLQREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGELLTVGSYDWAG 377
QY 361 GVFLYTSKSTFTINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHTGKAVFT 420
DB 378 GAFLYPPNMSPTFINMSQENVMDSDYLGYSYTELALWGVQSLVGLGAPRYOHTGKAVFT 437
QY 421 QNTGWMESNANVKTQIGAYFGASILCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 480
DB 438 QVSRQWRMKAETVGTQIGSYFGASILCSVDVDSNGSTDVLVIGAPHYEQTGQGVSVCP 497
QY 481 PRGQARWCCDVLVYGEQGWGFRGAALTVLGDVNGDKLTDVAIGAPGBENRGAVYLF 540
DB 498 PRGMR-RWMCDAVLYGEQGWGFRGAALTVLGDVNGDKLTDVAIGAPGBENRGAVYLF 556

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds

(without alignments)
7512.163 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5868	99.9	1153	1 RWU1B	cell surface glyco
2	4476	76.2	1153	2 S00551	leukocyte surface
3	3483	59.3	1163	1 RWU1C	cell surface glyco
4	1548.5	26.4	1170	2 S03108	cell surface glyco
5	1533.5	26.1	1163	2 I56126	lymphocyte fuction
6	1149	19.6	1179	2 A53213	integrin alpha-E c
7	1102.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1084	18.5	1170	2 I45914	integrin alpha 2 s
9	1072	18.2	1178	2 S44142	VLA-2 protein homo
10	1069	18.2	1161	2 A33998	integrin alpha-2 c
11	1060	18.0	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	633	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I59409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alaph integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subun
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	543.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	530.5	9.0	1072	2 A38457	integrin alpha-6 c
25	529.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1031	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.5	1146	2 S40311	integrin - fruit f

30 495.5 8.4 1039 2 A34269 integrin alpha-2b
31 494 8.4 1394 2 A29637 position-specific
32 490 8.3 1137 2 JC5950 integrin alpha-7 c
33 488 8.3 1135 2 I61186 alpha-7 integrin -
34 486 8.3 126 2 B30892 leukocyte adhesion
35 485.5 8.3 1037 2 A60163 glycoprotein IIB -
36 469.5 8.0 1106 2 S38783 integrin alpha cha
37 466 7.9 1226 2 S44924 F54F2.1 protein -
38 454 7.7 1045 2 S60571 integrin alpha v c
39 445.5 7.6 1139 2 S28277 hypothetical prote
40 422 7.2 1115 2 T09403 integrin alpha cha
41 417.5 7.1 1115 2 T09433 integrin alpha cha
42 391 6.7 764 2 I36916 glycoprotein IIB -
43 309 5.3 1086 2 T18523 integrin alpha cha
44 299 5.1 604 2 I36917 glycoprotein IIB -
45 296.5 5.0 272 2 A55348 integrin alpha-1 -

ALIGNMENTS

RESULT 1

RWU1B

cell surface glycoprotein CD11b precursor [validated] - human
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; 152567
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CR
B.
A:Reference number: A31108; MUID:88315033; PMID:2457584
A:Accession: A31108
A:Molecule type: mRNA
A:Residues: 1-1153 <COR>
A:CROSS-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A:Note: Part of this sequence was confirmed by protein sequencing
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor b
A:Reference number: A28915; MUID:88257215; PMID:2454931
A:Accession: A28915
A:Molecule type: mRNA
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A:CROSS-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
A:Note: the authors translated the codon TAC for residue 1129 as Thr
A:Note: part of this sequence, including the amino end of the mature protein, was confir
R:Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A:Reference number: A41600; MUID:92073318; PMID:1683702
A:Accession: A41600
A:Molecule type: DNA
A:Residues: 1-9 <SHE>
A:CROSS-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A:Reference number: A34193; MUID:88190151; PMID:2833753
A:Accession: A30892
A:Molecule type: mRNA
A:Residues: 917-1042 <AR2>
A:CROSS-references: GB:M18044
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A:Title: CDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
A:Reference number: A32218; MUID:89098893; PMID:2563162
A:Accession: A32218
A:Molecule type: mRNA
A:Residues: 9-1153 <HIC>
A:CROSS-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
 R.Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 490-490, 1993
 A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FLE>
 A:Cross-references: GB:S52227; NID:9263047; PIDN:AA24821.1; PID:9263049
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R.Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:35339202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R.Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A>Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: I52567; MUID:92144986; PMID:1346576
 A:Accession: I52567
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:9180184; PIDN:AA51960.1; PID:9553219
 A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: ITGAM; CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A>Note: promoter contains a GATA motif and two Spl consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homod
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAP>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:148-318/Region: von Willebrand factor type A repeat homology <VWA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding
 Query Match 99.9%; Score 5868; DB 1; Length 1153;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ENLDTENAMTFQENARGFGQSVVQLQSGRWVVGAPQEI VAAHQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRWVVGAPQEI VAAHQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSLNLRQPOK 120
 DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQCSENTYVKGCLFLFGSLNLRQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKPEVSTVMEQKKSKTLFSLMOYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKPEVSTVMEQKKSKTLFSLMOYSEEF 196
 QY 181 RIHFTKFEQNNPNRSLVKPIITOLLGRTHATGVKRVIRELLNITNGARKNAKILIVI 240
 DB 197 RIHFTKFEQNNPNRSLVKPIITOLLGRTHATGVKRVIRELLNITNGARKNAKILIVI 256
 QY 241 TDGKEGDPGLGYEDVTPRADREGVIRVIGVGDAFPRSEKSRQELNTIASKPRDHDVQYN 300
 DB 257 TDGKEGDPGLGYEDVTPRADREGVIRVIGVGDAFPRSEKSRQELNTIASKPRDHDVQYN 316

QY 301 NFPAALXIQNOLREKIFAIEGTQTGSSSSPEHEMSQSGFSAATTSNGPLLSSTVGSYDWAG 360
 DB 317 NFPAALXIQNOLREKIFAIEGTQTGSSSSPEHEMSQSGFSAATTSNGPLLSSTVGSYDWAG 376
 QY 361 GVFLYTSKESKSTFINMTTRVDSNDNDAYLGAAAIILNRRVQSVLGAAPRYQHIGLVAMFR 420
 DB 377 GVFLYTSKESKSTFINMTTRVDSNDNDAYLGAAAIILNRRVQSVLGAAPRYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVVCP 480
 DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVVCP 496
 QY 481 PRGQARWQCDVLYGSGQGFWRFGAALTVLGDVNGDKLTDVAIGAPGSEDNRGAVLYF 540
 DB 497 PRGQARWQCDVLYGSGQGFWRFGAALTVLGDVNGDKLTDVAIGAPGSEDNRGAVLYF 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMFNPFEVARNVFECDQVVKGEAGEVRVCLHVKQKSTRDLRREGQIQSVVT 660
 DB 617 PVLRVKAIMFNPFEVARNVFECDQVVKGEAGEVRVCLHVKQKSTRDLRREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPVSPILVRLNF 720
 DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPVSPILVRLNF 736
 QY 721 SLVGTPLSAFQNLPRVLAEDAQRLLFTALFFPKKNCNDNICQDDLSITFFSMSLDCLVVG 780
 DB 737 SLVGTPLSAFQNLPRVLAEDAQRLLFTALFFPKKNCNDNICQDDLSITFFSMSLDCLVVG 796
 QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLLACESASSTEV 840
 DB 797 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFENSESVTNIITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
 DB 857 SGALKSTSCSINHPIFENSESVTNIITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
 QY 901 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVQVSMGQSLPISLVFLVVP 960
 DB 917 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVQVSMGQSLPISLVFLVVP 976
 QY 961 RLNOTVIWDRPQVTFSENLSSTCHTKGRLPSSHDFLAELRKAPVVCNSIAVQRIQCDIP 1020
 DB 977 RLNOTVIWDRPQVTFSENLSSTCHTKGRLPSSHDFLAELRKAPVVCNSIAVQRIQCDIP 1036
 QY 1021 PFGIOEENFATLKGMLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTET 1080
 DB 1037 PFGIOEENFATLKGMLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTET 1096
 QY 1081 KVEPPEVNPPLIVGSSVGGGLLILALITATLALYKLGKPKQYKDMMSGGPPGAEPO 1137
 DB 1097 KVEPPEVNPPLIVGSSVGGGLLILALITATLALYKLGKPKQYKDMMSGGPPGAEPO 1153
 RESULT 2
 S00551
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N:Alternate names: complement-3 receptor alpha chain
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
 S:Accession: S00551; I59078
 S:Pyrela, R.
 EMBO J. 7, 1371-1378, 1988
 A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the i
 A:Reference number: S00551; MUID:88312584; PMID:3044779
 A:Accession: S00551
 A:Molecule type: DNA
 A:Residues: 1-1153 <PYT>
 A:Cross-references: EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:952983
 A>Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:G198993; PIDN:AAA39484.1; PID:G554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.2%; Score 4476; DB 2; Length 1153;
Best Local Similarity 74.3%; Pred. No. 1.2e-301;
Matches 845; Conservative 142; Mismatches 145; Indels 2; Gaps 2;

Qy 1 FNLDTENAMTFQENARGQSQVYVQLGSRVYVVGAPQEIIVAAANQKSLYQCDYSTGSCPEI 60
Ddb 17 FNLDTEHPMTFQENAKGFCQNVVLGSGTSVVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76

Qy 61 RLQVPVAVNMISLGLSLAATTSPOLLACGPTVHOTCSENTYVKGLCELFGSNLRQOPQK 120
Ddb 77 PLQVPPEAVNMISLGLSLAVTSPQOLLACGPTVHONCENTYVNGLCYDFGSLNLRPQQ 136

Qy 121 FPEALRGCPQSDIDAFIDGSGSIIPHDFRMKEFVSTWELKSKSLFSLMYSEEF 180
Ddb 137 FPEALRGCPQSDIDVFLIDGSGSINNIDFQKKEFVSTWELKSKSLFSLMYSEEF 196

Qy 181 RHFTKFEQNNPNRSLVKPTOLLGRTHATGKVKVIRELLNTNGARKNAFKILLVI 240
Ddb 197 RHFTFNDFKPNRSPRSHVSPKQLNGRTKTASGIRKVVRELPHKNGARENAKILLVI 256

Qy 241 TQGEKFGDPLGYEDVTPEDRGGVRYVIGVGDAPKSKSRQELNTIASKPRDHFVQVN 300
Ddb 257 TQGEKFGDPLGYEDVTPEDRAGVRYVIGVGNVFNKPSRRELDTIASKPAGEHVFQD 316

Qy 301 NFEALNTIQNLRKIFAIEGTQTSSEFHEMSQEGFSAITNGPLLTSTVGYDWAG 360
Ddb 317 NFEALNTIQNLRKIFAIEGTQTSSEFHEMSQEGFSAITNGPLLTSTVGYDWAG 376

Qy 361 GYFLYTSKESFTINATRYVDSQNDAYLGYAAITLNRVQSLVGLGAPYQHIGLVAMPR 420
Ddb 377 GAFLYTSKDKVFTINTIRVDSQNDAYLGYAAITLNRVQSLVGLGAPYQHIGLVAMPR 436

Qy 421 QNTGWMESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 480
Ddb 437 ENFGTWEPHTSIKSGIQSIFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 496

Qy 481 PRGQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDAVAGPGEEDNRCGAVLYP 540
Ddb 497 PRG-RARWOCEALLHGDQHPWGRFGAALTVLGDVNGDKLTDAVAGPGEEDNRCGAVLYP 555

Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGCAQGHVILLRSQ 600
Ddb 556 YGASIASLASHSRRIIGAHFSPGLQYFGQSLGGQDLTMDGLVLTGCAQGHVILLRSQ 615

Qy 601 PVLRVKALMEPNPREVARNVFNCDQVKGKAGVRCVCLVOKSTRDLREGQIOTSVTI 660
Ddb 616 PVLRLKALMEPNPREVARNVFNCDQVKGKAGVRCVCLVOKSTRDLREGQIOTSVTI 675

Qy 661 YDLALDSGRPHSAVFNETKNSRTROTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNF 720
Ddb 676 YDLALDPVRSIRIAPFDETKNTNTRRTQVGLMOKCETLKLILPDCVDVSPVILRLNY 735

Qy 721 SLVGTPLSAPCNLRPLVAEDAQLFTALPFKFNQGNINICODDLISITFSFMSLDCLVWG 780
Ddb 736 TLVGEPLRSFGLNRPVLAEDAQRFFTAAMPFEKFCGNDSICODDLISITFSFMSLDCLVWG 795

Qy 781 GPREFNVTVTRNDGDSVHTQVTFPPPLDLSTVRKVTSTLQVRSQSRWRL-ACESASSTE 839
Ddb 796 GPQDFNMSVTLRNDGDSYGTQVTVVYPSGLSRKDSASQNPITKKFPWVKPAPSSSSSE 855

Qy 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKYTSNNMPTKTE 899
Ddb 856 GHGALKSTTWNINHPFPANSEVTFNITFDVDSKASLGNKLLKANKYTSNNMPTKTE 915

Qy 900 FOELPVPKAYVMVTVSHGVSTKYLNFASNTSVNMQHOYQVSNLCQSLPISLVLVP 959
Ddb 916 FOELPVPKAYVMVTVSHGVSTKYLNFASNTSVNMQHOYQVSNLCQSLPISLVLVP 975

Qy 960 VRLNQTVIDRPOVTVSENLSSCTCHTERLPSSDPLAELRKAFVNCSTAVCORICODI 1019
Ddb 976 VQINNVTVDHPQVIFQNLSSACHTEQKSPFNSHNFQDLERFVLCNSVAVCKRIODL 1035

Qy 1020 PFGIOQEFNATLKNLSFDWYINTKSNHLLIVSTABILFNDVSTLPLPGCAFVRSQTE 1079
Ddb 1036 PFPNTQIEFNATLKNLSFDWYINTKSNHLLIVSTABILFNDVSTLPLPGCAFVRSQTE 1095

Qy 1080 TKVEPFPVPLPLIVGSSVGGILLALITAAALYKLGFPKQYKDMSEGPPCAEPQ 1137
Ddb 1096 TKVEPFPVPLPLIVGSSVGGILLALITAGLYKLGFPKQYKDMSEGPPCAEPQ 1153

RESULT 3
RHHUIC
cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999
C:Accession: A36584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Note: This revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: This sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M16195; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on m
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: calcium, cell adhesion; glycoprotein, heterodimer; magnesium; tandem repeat
F:1-19/Domain: signal sequence status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pr

A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBIP:124326)
F:142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 9.3e-66;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

QY 1 FNLDENAMTFOENARG-FGQSVVQL---QGSRVVVGAPQEIIVAAORGSLYQCDYGTGS 56
DB 1 FNVVKNMTFSGPVEDMFGVTVQYENEBGKWLLGSLPLVGQPKNRTGDKVYKCFVGRGE 60

QY 57 CEP-IRLQVPEA-----VMSGLSLAATSPOLLACGPTVHQTCSNTYVKGL 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKNMTFGSTL-VTNPNNGFLACGFLVAYRCGHLHYTGI 119

QY 107 CFLFGSNLRQOPQKPPFALRCPOEDSDIAPLDGSGSIIPHDPRRMKEFVSTVMEQLK- 165
DB 120 CSDVSPFFQVNSIAP-VQECSTQ-LDVIIVLQSGNSIYPWD--SVTAFNLDLKEMDI 174

QY 166 -KSTLFLSLMOYSEEFRIHFTFKFQNNPNRSLVKPIITQLLGR-THTATGVRKVIKRELL 223
DB 175 GPKQTOVGIVQYGENVTHEFNLYKYSSTEEVLVAAKKIVQGRGRTWTALGTDTRKEAF 234

QY 224 NITGARKNAFKILIVTDGKFGDPLGYEDVIPEADREGVIRVIVGVDAFR-----SE 278
DB 235 TEARGARGVKVWVITVDGESH-DNRLKVKVIOCEDENIQRESIALGSYNREGLSTE 293

QY 279 KSRQELANTIASKPPDRHVFQVNNFEALKTIQNLREKIPAEIGTQTSSEFHEHMSQEG 338
DB 294 KFVEBIKSIASEPTEKHPFNVSDELAVITVKTGERIFALEATADQSAASFEMENSQTG 353

QY 339 FSAITSNGLPSTLTVGSDVAGVPLYSKE-----KSTP-INNVVDSMDNDVLAGVAA 392
DB 354 FSAHYSQDWMVLGAVGADVNGTVVMQKASQIIPRNTFVNSTKKNELP-ASTLGTIV 412

QY 393 AILLENRVQSL-VLGAPRYQHIGLVAMPRQNTGHWESNANVKGTOIGAYFOASLCSVDVD 451
DB 413 NSATASSGDVLYIAGQPRYNTGQVILVYRMEDGNKIKILQTLGSGEIGSYFGSILTTDID 472

QY 452 SNGSTDLVLGAPHY-----YEQTR-GQVSVCLPGRQARWQCDVLY 495
DB 473 KDSNTDILLVAPMYMGTKEBEQGVVYVYALNTRFQMSLEPIKQTCSSRQHNQCTT 532

QY 496 GEQGPWG-RFCALTLVLDGVNGDKLTDVAGPEEDNMGAVLYLFGTSGSGISPSHSQ 554
DB 533 ENKHEPCGARFGTAAVVDLNLGDFNDIVIGAPLEDHGGAVIYHG-SGKTIRKEYAQ 591

QY 555 RIAGKLSPLQYFQOGLSGGODLTMDGLVDLTVAQGHVLLRSQFVLVRKAIWENFR 614
DB 592 RIPSQGGDKTLKFQOISIHGMDLNGDGLTDVTIGLGAALFWSRDVAVVYKVTMNPEN 651

QY 615 EVARNVFCNDQVVKGEAG-EVRVCLHVQ-KSTRDELREGQISQSVVYDLALDSGRPH 671
DB 652 KVNIOKKNCH--MEGKETVCINATVCEVKLSKEDTIYRADLQ-----YRVTLDSLRQI 704

QY 672 SRAVFNET-----KNSTRRTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNPLSVGT 725
DB 705 SRSFSGTQERKQVQNRNIVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLTD 755

QY 726 PLSAFNGLRPVLAEDAQLFTALFPFKNCGNDNIQDGLSITFSFMSLDCLVVGPRE- 784
DB 756 PENG-----PVLDDSLPNSVHEIYFPKXDCGKKEKISDLSLHVAETTEKULLIVRSQDK 810

QY 785 FNVTVTRNDGDSYRTQVTFPPFDLSYRVKYSTIQNORSQSWELACESASSTEVSGAL 844
DB 811 FNVSLTVKNTKDSANTRTIVHYSNLFVSGIEALQKD-----SCSEN----- 853

QY 845 KSTCSINHPFPENSEVTFNITFDVDSKASLGN-KLLKKNVTSNNMPTNKTEQLE 903

RESULT 8

I45914
Integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin
A:Reference number: A54402; PMID:94193647; PMID:7511592
A:Accession: I45914
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: GB:L25886; NID:9439695; PIDN:BA859255.1; PID:9439696
F:161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 1.8e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDENAMTFO-ENARGGQSVVQL---QGSRVVVGAPQEIIVAAORGSLYQCDYGTGS 54
DB 19 YNVGLPKAKITFGSFSBQGYAVQGFIPNKGNNLLVGSWGFPRKRMGVDVYKCPVDLST 78

QY 55 GSCBPIRLQ-----VPVEAVNWSLGLSLAATSPOLLACGPTVHQTCSNTYVKGLC 107
DB 79 TTCEKLNQTSMSNVTEMKTNWSLGLTLRVNVTGGFLTCGPLMAQCCSQSVYTTGVC 138

QY 108 FLFGSNLRQOPQKPPFALRCPOEDSDIAPLDGSGSIIPHDPRRMKEFVSTVMEQLK-- 165
DB 139 SDVSPDP-QLTSTFAPAVQTCF-SFIDVVVVCDESNSIYPWD--AVKNFLEKFPVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLVKPIITQLL-----GRHTATGVRKVIKRE 221
DB 195 PTKTQMLQIYANNPRVFNLTNFKSD---EMKATSTQTFQYGGDLTNFKAIQVARDT 251

QY 222 LNMITNGARKNAFKILIVTDGKFGDPLGYEDVIPEADREGVIRVIVG-----GDAPR 276
DB 252 AYSTAAGRGPRGATKVMVWVTDGESH-DGSKLKAVIDQCNKDNILRFGIAGVLGYNLRNLD 310

QY 277 SEKSRQELANTIASKPPDRHVFQVNNFEALKTIQNLREKIPAEIGTQTSSEFHEHMSQ 336
DB 311 TKQLIKELKAIASIPTERHFFNVSDDEADLLEKAGTIGEIQIFSGIEGTQVQ-GDNFQEMSQ 369

QY 337 EGFSAAIT--SNGPLSTVGSYDAGVFLYTSKESKSTFTNMT--RVDSDMN-DAYLGYA 391
DB 370 VGSFSAEYSPQNTILMLGAVGAYDSGTVQVTKPHGLIFSKQAEQIQLDRNHSSYLGYS 429


```

Qy 956 FLV-----PVLNQTVDWRPQVTF-SENLS 980
D5 985 LVTHIPQYTKENPLLYLTGQTQDAGDISCTAEINPLKLPHTA-----DSVSFKNENR 1040
Qy 981 STCHTKERLPSHSDFLAELKAPVWNSIAVCQIQCDIPFGIOEBRNATLKGNSLFDW 1040
D5 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHKAKBYFINVTVRWNR 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDSTFLAPGQAFVRSOTETKVEFFVFNPLPLIVGSSVG 1100
D5 1081 FAASTFQTVQLTAAEIDTHNPQLFVIEENAVTIPLMIMKPKERAEVPT--GVIIGSIIA 1138
Qy 1101 GLILLALITAAALYKLGFFKQYKOK 1125
D5 1139 GILLALLAMTAGLWKLGFKQYKOK 1163

RESULT 10
A:33998
Integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAX>
A:Cross-references: GB:X17033; MID:g33906; PIDN:CAA34894.1; PID:g33907
A:Note: the authors translated the codon CAT for residue 802 as Gln, GTC for residue 803
R:Catimel, B.; Parentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIIb and GPIIc)
A:Reference number: A56793; MUID:92061944; PMID:1963640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements.
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:L24121; MID:g400342; PIDN:AAA16619.2; PID:g4583535
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 2e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

Qy 1 ENLDTENAMTQ-ENARFGGSVQL-----QSSRVVVGAPQIVAAQNGSLYQC--DYST 54
D5 30 YNVGLPEAKIFGSPSSFOGYAVQGFNPKGNWLLVGSFGFPENRGDVKCPVDLST 89
Qy 55 GSCPEIRLQ-----VPVEAVNMELGLSLAATTSPPOLLACGPTVHQTCSNTYVKGLC 107
D5 90 ATCEKLNLTQSTSIPNVTEKTNMSLGLILTRNMTGGLGCLPWAQCCNQYVTTGVC 149

```

```

Qy 108 FLFGSNLRQOPKQFPALRGCCPOEDSIAPLIDGSGIIPHDPRMKBFVSTVMEQLK-- 165
D5 150 SDISPDV-QLSASFSPATQPCSL-IDVVVCCDESNSIYPWD--AVKNFLKFTVQGLDIG 205
Qy 166 KSTLSLMOYSEBFIHFTFKFQNNPNRSLVKDITQLLG-RTHTATGVKRVIRELN 224
D5 206 FTKTQVGLQIYANNRVVFNLTITKTKEMIVATSTQSYGGDLTNTFGAIQVARKAYS 265
Qy 225 ITNGARKNAFKILIVITDGEKFDPLGYEDVITPEADREGVIRYVIGV-----GDAPRSEK 279
D5 266 AASGGRSRTKVMVVVTDGESH-DGSNMLKAVIDQCNDHNDILRPGIAVLGYLNEALDTKN 324
Qy 280 SRQELNTIASKPRDRVFNQNNFEALKTQIQRIKIPALECTQVSSSSFEHEMSQEGF 339
D5 325 LIKEIKATIASIPTRYFFNVSDAALLEXAGTGEQIFSEIEGTVOG-GDNFQEMSVQGF 383
Qy 340 SAAITNGP--LLSTVGSYDWAGGVFLYTSKSKSTTINMT--RVDSDMN-DAYLGAAAI 394
D5 384 SADYSSQNDILMLGAVGAFMSGTIVQKTSKSHLIFPKQAFDQILODRNHSSVLYGVAA 443
Qy 395 ILNRVQSVLGAAPRQOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFASLCSV 448
D5 444 ISTGSESTHFVAGAFRANVTQIIVLYSVN-----ENGNTITVIAHRGDQIGSYFGLCSV 498
Qy 449 DVDSNGSTDLVLGAPHYEYQTR--GQSVSVCLPRGQRARWOCDAVLYGEOQOPWGRPG 506
D5 499 DVDKDTITDVLVGAPWMSDLKKEGRVYLFPIKGIILQHQ---FLEGPGIENTRPG 555
Qy 507 AALTVLGDVNGDKLTDVAIGAPCEEDNRGAVLYPHGTSGSGISPSHSQRIAGS--KLSPR 564
D5 556 SAJAAALSDINMGDFNDVIVGSPLENQSGAVIYNGHQGT-IRTKYSOKILGSDGAFRSH 614
Qy 565 LQYFGSLSGGQDLTDGLVLTVGQGHVLLRSQPLVLAIRMBNPREVARNVPECN 624
D5 615 LQYFGRSLQYGLNGDSITDVSIGAPGVVQWMSQSIADVAIEASTPESKI--TLVNKN 672
Qy 625 DQVVGKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALD---SGRPHSRVAFNETK 680
D5 673 AQII-----LKLCP-----SAKERPTKQNNVAIVNITLDADGFSRVTSGRLPKENN 721
Qy 681 NSTRQTVGLTQTC--ETLKLQNCIEDPSPVILRNPSILVGTPLPAGNLRVLA 738
D5 722 ERLQKNMVMVQACSPHEIYIQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFPPEKNCNDNICDDLSITF---SFMSLDCLVGVGPFNFVTVVRND 794
D5 773 AYSETAKVFSIPFAKCGEGDGLCISDLVDVQIPRAEQEPFIVSNQKELTSTVLKNK 832
Qy 795 GDSYRQVTFPPLDLVYKVKSTLQNRQSRWRLACESAST-EVSGALKSTSCSINH 853
D5 833 REGAVNTGIVVDFSENLF-----ASFSLPVDGTGTEVTCQVAASQKSVACDVGY 880
Qy 854 PIPENSEVTNLTEDVDKASLGNKLLKANVTSENNMPRTNKTETFOLELPVKIAYMV 913
D5 881 PALKREGQVTFITNDFNLO-NLQNOASLSFQALSQSEKADNLVNLKIPPLYDAEI- 938
Qy 914 VTSHGVSTKYLNFTASENTSRVWQHYQVSNLQOR-----SLPISLVFLV----- 958
D5 939 ---HLRSTNINFEIISDGNVPSIVHSFEDVGPKEIFSLKVTGTGSPVSMATVHIHPQ 995
Qy 959 -----PVLNQTVIDWRPQVTF-SENLSSTCHTKER 988
D5 996 YTEKKNPLMYLTGVQTDKAGDISCNADINPLKIGQT-----SSSVSPKSENFN---HTKE- 1047
Qy 989 LPGHSDFLAELKAPVWNSIAVCQIQCDIPFGIOEBRNATLKGNSLFDWIKTSHNH 1048
D5 1048 -----LNCRTASCSNVTCWLDKVMHMKGEYFVNVTVTRWNGTFASTFQT 1091
Qy 1049 LLIVSTRABI-LFNDSTVFTLLPGQAFVRSQTEKTVKPFVEFNP-----LP--LIVGSSVG 1100
D5 1092 VQLTAAAEINTNPBIYVI-----BDNTVTIPLMIMKPKERAEVPTGVLIIGSIIA 1141

```

[illegible]

RESULT 12

A41131

lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N;Alternate names: integrin alpha-4

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999

C;Accession: A41131; S16742

R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.

J. Cell Biol. 115, 1149-1158, 1991

A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte-

A;Reference number: A41131; MUID:92064645; PMID:1840602

A;Accession: A41131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1039 <NEU>

A;Cross-references: EMBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51485

C;Superfamily: integrin alpha-4 chain

C;Keywords: cytoskeleton; transmembrane protein

```

Query Match      11.3%; Score 663; DB 2; Length 1039;
Best Local Similarity 22.7%; Pred. No. 2.3e-37;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

Qy      1  FNLDTENAMTFQ-ENARGGQSIV-QLQSR--VVGAQOEIVAA----QBSGLQQCDY 52
      :||| :|: :|: ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db      41  YNLDPENALLYQPSGTGLFGYSVVLTHSGSKRWLIIVGAPTASWLSSAVVNPGRTRCGI 100

Qy      53  STG---SCEPIRLQVP-----VEAVNLSLGLSLAATTSP-POLLAQG---PTVHQ 95
      :||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db      101  RKPNQTCQLQSGSPSPGCKGTCLERENQWLVTLSRQPGENGSI VTCGHRWNIFY 160

Qy      96  TCSENTYVVKGLCPLFGSNLRQOPKQPEALRGCPQSDIAFLIDGSGSI PHDFRRMKE 155
      :||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db      161  MKSDNKLPTGICVYMSDSDRULSK----- 185

```

```
Qy 156 FVSTVMEQLKSKTFLSLMOYSEEFRIHFTKFEQNNPNRSLVKPITOLLGRHTATGV 215
Db 186 -----RWAPCVKDYT----- 195
Qy 216 RKTRELLNITNGARKAKFLIVITGEKEGDLGVEDVPEADRGVIRYVIGVDAP 275
Db 196 -----RKPE----- 200
Qy 276 RSEKSRQELNTIASKPPRHHVFOVQNNFEALNTIQOLREKIFALEGQTQSGSSSFEHMS 335
Db 201 -----NFAS-----C 205
Qy 336 QEGFSAITNSGPNLLSTVGSVDNAGGVPLY--TSKEKSTFINMTVRVSDMNDAYLGYA-- 391
Db 206 QAGISSFYTDLIYMGAPGSYMTGTTFVYNNITNQYKAFVD--RQNVKFGSGLVSVG 263
Qy 392 AAIIILNRVQSLVGLAPRYOHIGLVAMFRONTGHWESNANV---KGTQIGAYFGASLCS 447
Db 264 AGHFRSHPTTEVVGAPQHEQIKAYIF---SIDENELNIYEMKGLGSLYFGASVCA 319
Qy 448 VVDNSGSTDVLVLGAPHYBQTRGGQVSVCLPRGQRA--RWQCDVILYGEQGFQWGRFG 506
Db 320 VDLNADGPSDL-LVGAPMQSTIREGRVFVY--INSGMGAVMVENRVLVGSDDKYA-ARFG 376
Qy 507 AALITVLGVNCDKLTDAIGAIPGEDNMGAVYLFHGTSGSISPSHSORTAGSKLSPLQ 566
Db 377 ESIALNGDIDNDGPDIAIGAPQDDLRGAVIITNGRV-DGISSTYSORIEGQISLSLR 435
Qy 567 YFGQSLSGQDLTMDGLVDLTVGA--QGHVILLRSQPLVRKAIMENPREVARNVFECD 624
Db 436 MFGQSIQSQIDADNNGYVDVAVGAPQSDSAVLLATRPVIVASLS--HPESVNETKFDCT 494
Qy 625 DOVVKGEAGEVRLVCLVOKSTRDLREGQIQSVVYDLDLDSGR---PHSEAYF--NET 679
Db 495 -----ENGLPSVCMELTLCFSYKGEVPGYIVLFYNVSDVHRKAEPSRFTFFNGT 547
Qy 680 KNSTRRTQVGLGTQCTETLKLQPNCTEDPVSPTVLRNPLS-----VGTFPSAFGNLR 734
Db 548 SDVITGSRVSSGSEKCTHOAFKQVDRDILTPHVEATVHLGHVITKNTSEEFPLQ 607
Qy 735 PVLAEDAQR-LFTALFPPEKNCNDNICQDLSITFS-----FMSLDCLVGGPRFNV 787
Db 608 PILQOKKEQVIRKMINFAFCAYEN-CSADLOVSARKVGLPKPYENKTYLAVGSMKTMIL 666
Qy 788 TVTVNDEGDSVKTQVTFPPFLDLSYKVSFLQORSORSLACESASTEVSGALKST 847
Db 667 NVSLFNAGDATTETLNVQLPTGLYFKILDLSEK-----QINCE---VTIESSGVK-L 716
Qy 848 SCISNHPFPENSEVTFNITFDVDSKASLGNKLLKANKVTSEN--NMPTKTKTFQLELP 905
Db 717 ACSLGYIVDRLSRIDISFLDVSLSRAHEDLSISVHASCEGELDQVRDNRVLTIP 776
Qy 906 VKYATVMTSHGV--STKYLNFPTASENTSRVMQHQ-----YQVSNLQORSLP-LSLVFL 957
Db 777 LREYEV--MLTVHGLVNPFSFVYGSEENEPETCMAEKLNLFTHVINTQISVAPNVSKIM 834
Qy 958 VEVRLNQTVINDRP--QVTFSENLSSTCKTE-----RLPSHSPFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLVDQTTTGQCHFGYHGRECTFAQCKGIAGTLTDIVKFLSKTD 891
Qy 1001 KAPVNVCSIA--VCORIQCDDIPFGIQEFPATLKNLSFMYIKTSHNLLIIVSTAEIL 1058
Db 892 KR-LLYCKMADQECFLDPLCN---FGKMSG-----KEASVHTQLGERPSIL 933
Qy 1059 FNDVSFTLLPGGQAFVRSTQTEKVPFFVFPNP----- 1090
Db 934 ENDETSSL-----KFEIKATAFPEPHKPVIELNKNDENVAHVLEGLHQRKPEHF 983
Qy 1091 -LPLIVGSSVGLLLALITAAIYKLGFPKQYKDMSE 1128
Db 984 TIIITISILLGLIIVLLISCVMWKAGFTRQYKSILOE 1022
```

```
RESULT 13
S06046
integrin alpha-4 chain precursor - human
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A;Molecule type: mRNA
A;Residues: 1-1038 <TAK>
A;Cross-references: GB:X16983; EMBL:X15356; NID:G33945; PIDN:CAA34652.1; PID:G33946
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A;Title: Characterization of the alpha-4 integrin gene promoter.
A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A;Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molec
A;Reference number: A94151; MUID:87264112; PMID:3033641
A;Accession: D28018
A;Molecule type: protein
A;Residues: 40-50, 'E', 52-53 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:I128032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transme
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 633; DB 2; Length 1036;
Best Local Similarity 25.2%; Pred. No. 2.8e-35;
Matches 248; Conservative 155; Mismatches 370; Indels 212; Gaps 39;

Qy 272 GDAPRSEKSRQELNTIASKPPR-----FAEGTQSGSSSFEHMSQEGFSAITNSGPLLSTVGSY 356
Db 121 GKTCLERDQNLGVLTSRQKENGSIIVTCGHRWKNIIFYIKNKENKLPCTGCGYVPPDLRT 180
Qy 311 QLEKEL-----FAEGTQSGSSSFEHMSQEGFSAITNSGPLLSTVGSY 356
Db 181 ELSKRIAPCYQDVYKXFGENPA-----SCQAGISSFYTKDLIVMGAPGSS 225
Qy 357 DWAGGVPLY---TSKEKSTFINMTVRVSDMNDAYLGYA--AAIILNRVQSLVGLAPRYQ 411
Db 226 YWTGSLFVYNTITNKYKAFDKQNVKF---GSLYGVSVGAGHFRSQHTTTEVVGAPQHE 282
Qy 412 HIGLVAMFRONTGHWESNANV---KGTQIGAYFGASLCSVDVSDNSGSTDVLVLGAPHY 467
Db 283 QIGKAYIF---SIDEKELNIHEMKKGLGSLYFGASVCAVDLNADGPSDL-LVGAPMQS 337
Qy 468 EOTRGQVSVCLPRGQRAWQC--DAVLYGEQGFQWGRFGAALTVLGDVNGDKLTDTVAIG 526
Db 338 TIREGRVFVY--INSGMGAVMVENRVLVGSDDKYA-ARFGESIVNLGDIDNDGPDVAIG 395
Qy 527 APGEDNRGAVYLFHGTSGSISPSHSORTAGSKLSPLQYFGQSLSGQDLTMDGLVDL 586
Db 396 APQDDLLQGAIIYNGRA-DGISSTYSORIEGQISLSMFGQSIISGQIDADNNGYVDV 454
Qy 587 TVGA--QGHVILLRSQPLVRKAIMENPREVARNVFECDNVVKGGEAGEVRLVCLHVQK 644
Db 455 AVGAFRSDSAVLLATRPVIVASLS--HPESVNETKFDCT-----VENGWPSVICDILT 506
Qy 645 STRDLREGQIQSVVYDLDLDSGR-----PHSEAYFNETKNTSTRQTVLGLTQTCETL 699
```



```

Db 507 CFSYKGVPGVIVLFFNMSLDVARKAESPFRFPYSSNGTSDVITGSIQVSSRANCRTH 566
QY 700 KLQLENCIEDPSPVIVLKLNSLVGTPLS-----AFGNLRPVLAEDAQR-LFTALFPFBK 753
Db 567 QAFMRKDVRLDITPQIEBAAYHLGPHVISKSTBTEFPPLQILOQKKEKIMKTIINFAR 626
QY 754 NCGNDNI CODDLSIT--PSEM-----SLDCLVVGPREFNVTVTVRNDGEDSYRQVTFPP 807
Db 627 PCAHEN-CSADLOVSAGIKPLPHENKTYLAVGSKTLMVNSLNFNAGDDAYETTLHVKL 685
QY 808 PLDSLYRVKSTLQNRORSRLACESASSTEVSGALAKSTSCSNHPIPPENSEVTFNIT 867
Db 686 PVGLYFIKILEEER-----QINCE---VTDNSGVQ-LDCSIGIYVDHLSDISL 735
QY 868 FDVDSKASLGNKLLKANVTSEN--NMPTNKTEPQLPQLPVKYAVVMVVTSHGVSTKLN 925
Db 736 LDVSSLSABEDLSITVHATENBEEMDNLKHRSVTVPAIPUKYEVKLTVHGVNPTSFY 795
QY 926 FTASNTSRV-----MORQYQVSNLQORSIP--ISLVFLVPVRLNQTIVMDRPQVTFSEN 979
Db 796 GSDNEPEBTCVBERKNMLTFHVTNGSMAPNVSVIIMPNSFS-----PQTDKLPNI 848
QY 980 -----STCHTKERLPSPHSLAEALRKAPVNCVSIACQRIQCDIPFFGLOEFNATLK 1033
Db 849 LDVQTTGECFENYQ-----RVCALEQ-----OKSAMQTLK 880
QY 1034 GNLSPDWIKTSHNHLIVSTAE-----ILFN-----DSVFTLLPGQGAFFVRSQTE 1079
Db 881 GIVRF--LSKTDKELLYCIKADPHCLNFCNFGWESGKSAVHIQLEGRPS-ILEMDE 936
QY 1080 TKVEFEV-----PNP-----LPLVGSVGGILL 1103
Db 937 TSALKFEIRATGFPFENPRVIELNKDENVAHVLEGLHHQRPKRYFTTIVISSLLGLI 996
QY 1104 LIALITAAALYKLGPFKROVKDMSE 1128
Db 997 VLLLSYVMWAKGFFKQYKSLQEB 1021

RESULT 14
I58409
Integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
A:Cross-references: GB:D25303; NID:9464180; PID:BAA04984.1; PID:9533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:I24158
C:Superfamily: Integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 4.5e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

QY 331 EHEMSQEGFSAALTSNGLSTVGSYDMAGGVFLTSKEKSTFINNT-RVDSNDMDAYIG 389

```

```

Db 190 EHSQCAGIAGFTBELVWGAFGSYNAGTIKVLNLT-DNTYKLNDENVIMNRYTILG 248
QY 390 YA-AAIILNVRVOSLVGLAPRYOHIGLVAMFR--QNTGMWESNANVKGQIQI GAYFGASL 445
Db 249 YAVTAGHFSPSTIDVVGGAPOQKIGKVIYFRADRRSGTLIKI FQASGKMGSYFGSSL 308
QY 446 CSVDVDSNGSTDLVLGAPHYHYSOTRGQVSVCPFRGQARWQCDVAVLYGEOQPMGRF 505
Db 309 CAVDLNGDGLSDL-LVGAPMFSIRDEGQVTVY-INRNGCALEE-QLALTGD-GAYNAHF 364
QY 506 GAALTVDGVNGDKLTDVAIGAFGEDNRAVYLPHGTSGSGISPSHSQRIAGSKLS PRL 565
Db 365 GESIASLDDLDNDFPDVAIGAPKEDDFAGAVIYHGDAG-GIVPOYSMKLSGQKINPVL 423
QY 566 QYQGSLSGGQDLTMDGLVDLTGA--QGHVLLRSQFVLVRKAIMFENPREVARNVFEC 623
Db 424 RMFGQSISSGIDMDGNGYPDVTVGAPMSDSVILLRARPVITVD-VSIFLPGSINITAPQC 482
QY 624 NDQVVRKEAGEVRV-CLHYQ-----KSTRDLREGQIQSVVTVYDLDALDSGRPHRAV 675
Db 483 HD-----GQOPVNCVLTTCFSPHGKHVPBEI---GLNYVLMADVAKKRGQMPRV 531
QY 676 F--NETKNSTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRNFSI---VGTPLS 728
Db 532 FVLGETMGQVTEKLIQTMYNEESTCRHYVAHKRRVQDVISPIVFEAAISLSEHVTGEER 591
QY 729 AFGNLRPVL-----AEDAQLFTALPFFKXNCGNDNI CODDLSITFSFM--SLD-----CL 777
Db 592 BLPLTLPVLRWKKGOKIAQKNQTV---FERNCRSED-CAADLQLOGLKLLSMDKLTLYL 647
QY 778 VVGPREFNVTVVRNDGEDSYRQVTFPPFDLSYRKVSTLQNRQSRWKLACESASS 837
Db 648 ALGAVRNISLNISSNLGDDAYDANVSNVSRLEFF-----INMOKEMGISCLELLES 701
QY 838 TEVSGALKSTSCSNHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
Db 702 DFL-----KCSVGPFPFMSKSKYBSFVTFDTSLS--GEEVLSPIVTAQSG--NTER 750
QY 898 TE-----FOLELPVKYAVVMVTS-----HGVTKYLNFTASENVS---RVMOHQY 940
Db 751 SESLHNDTLVLPVLAHEVDTSITGIMSPTS FVYGESVDAANFIQLDLECHFPINITL 810
QY 941 QVSNLQORSIPISLVFL-VPVRLN-----QTVIWDPRQVTFSENLSSTCHTKER 988
Db 811 QVNTVGTSTLPGSVSISFNNRLSSGGAEMFHVQEMVVGQEKNCFSQKQFTPCIIIPQEQ 870
QY 989 LPSSHDFLABLRKA-----PVNCSIAVCQRIQCDIPFFGQIEEFNATLKGNLSFD 1039
Db 871 ENIFHTTFAFTKSGRKVLDCRKPGISCLTAHCN-----FSALAKESRTI-----D 917
QY 1040 WYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFFVRSQTSOTETKVEP----- 1084
Db 918 IY-----MLLNT-SILKDDSSVIQ-----FMSRAKVVDPPALRVVLEIAHNPVEV 962
QY 1085 ---FEVFN---PLPLIYV-----SSVGGLLLLALITAAALYKLGFFKQYKDM 1126
Db 963 TVVFEALHNLEPRGVVVGWIIAISLLVGLIFILLAVLLWKGFFRRRYKXII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

A;Residues: 1-1041 <HER>
A;Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:ARC23572.1
A;Experimental source: developmental stage embryo
C;Function:
A;Description: binds laminin
C;Superfamily: integrin alpha-2b chain

Query Match 10.58; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.88; Pred. No. 5.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY	336	QEGFSAITSGNP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS----
DB	181	QAGFSGIIFSDNSALVNGAGSYVLOQIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
QY	390	YAAAI--ILNRVQSLVLCAPRYQHI-GLVAMPQNTGMWESNANVKTQTOIGAYFGASLC 446
DB	240	YSLALGDFNGDGVDYVYVGTFRASLMLGLVAIFDQNLNQFN---QVMGTQIVAFYGSVT 296
QY	447	SVYVDSNGSTDLVLIGAPHYEYQTRGGQVSVCLPRGORARWQCDVLYGEQ----- 498
DB	297	VVDI--NNDTYDLLVAGFMYMDGPAIQ-----RWEGAVVYVLQNPDVGPGA 343
QY	499	-----GQPMWGFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546
DB	344	SNRLSLSTLIGGQTRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYIYHG--SAN 402
QY	547	GISPSHSQRIAGKLS--PRLQYFGOSLGGQDLTMDGLVDLTVGQA--GHVLLRSQPV 603
DB	403	GLKSTPAQVLFTSTLGHSGITTFGFSLOGGQMDKXKYPDLLVGAESANTAVLIRTPVV 462
QY	604	RVKAIMFNPREVARNVPECNDQVVYKGEAGVRVCLHVQKSTRDRLRREGIQSV----- 658
DB	463	SILDAFLNTEP-----IGINLENKTYE--LADGTMVTSFIAMT 497
QY	659	-----VTYDLALDSG--RPHSRAVENETKNSTRQTOVGLL-TOTCETLKIQ 702
DB	498	CFTYTGNYLPDHIDISYTVTVDSGIIANRRAMFVNDMSEITKRRRLAVSTQFCDFLEAY 557
QY	703	LPNCIEDVPSVILRLNFSVLGTPLSAFGN-----LRPVLAEADAQRLFTALFPPEK 753
DB	558	VGNSIEDKLTPIKVTLOYDL-----NNDSSRLQPHLEILLPIDMATMSTQTKQVSIQN 609
QY	754	NCGNNDICODLSITFSFMSGLDCLVVGPREPNVTTVNRNDEGDSYRTQVTFPFELDLSY 813
DB	610	NCVN--NICIPDLDTVT--PNLEPNIVIGQTELTQVSLNNGEDAFQSSLSVYIPGLQF 667
QY	814	RKVSTLQNRQSQRSLACESASSTEVSCALKSTCSINHPIFPEN-----SEVTFNIT 867
DB	668	VRL-----ERKAMDFTSVTCESD-----LRITCDTGNPMVKNNILFGLTLSTFOVS 717
QY	868	FDVDS-----KASLGNKLLKANVTSENMPRTNKTPEQLPVRKAVYVMVTSHGVS TK 922
DB	718	GDKDSIEFVYKAESENS--EDPNTLENNELNMTVPVTVDTLTKLSASYPEIWNYSTOED 775
QY	923	YL-----NFTASENTSRVMOHQYV-----SNLQSRSLPIS-----LVFLVPV 960
DB	776	YVVPFPKXNASEADIGMEVM-HLYEVNRTGSGNAGEVSLNQMPQKNEDGEYLFYLLGI 834
QY	961	RLNQTVIMDRPQ-----VTFSENLSSTCHTKERLPSHSDFLAELKAPVV 1005
DB	835	MTEEGVTCQLTQKANKPBGVKLEPSTKAKLSNSTQVSGRKRREPEVAEALACTDN--VI 892
QY	1006	NCSIAVCQRIQCDIPFGIQEENAT-----LKNLSFDWYIKTSHNHLIVSTAELP 1059
DB	893	YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY	1060	NDSVFTLLPGQGFVRSQETKVE--PFVFPNP-----LP----- 1092
DB	929	QKAVSELTPVQVQATIASAAVKTIPIYNIPLPRDFSDSTKASTLVTTTELVPPTPIAM 988
QY	1093	-LIVGSSVGLLLALITALYKLGPFKR-----QYKDMMS--EGGPP 1132

DB 989 WIIVSVLGGIILLIILGLWKCGFFERKXPGEEKYAPVASADKGGPP 1038
Search completed: June 7, 2004, 17:18:19
Job time : 25.559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds
(without alignments)
5937.039 Million cell updates/sec
Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENMTFQENARGFGQ.....FKRQYKDMSEGPPGAEFQ 1137
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42:*

ALIGNMENTS

RESULT 1		ITAM_HUMAN		STANDARD;		PRT; 1152 AA.	
ID	ITAM_HUMAN						
AC	P11215;						
DT	01-JUL-1989 (Rel. 11, Created)						
DT	01-OCT-1996 (Rel. 34, Last sequence update)						
DT	10-OCT-2003 (Rel. 42, Last annotation update)						
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)						
DE	(Neutrophil adherence receptor).						
GN	ITGAM OR CR3A OR CD11B.						
OS	Homo sapiens (Human)						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
OX	NCBI_TaxID=9606;						
[1]	SEQUENCE FROM N.A.						
RP	MEDLINE=88315033; PubMed=2457584;						
RX	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;						
RA	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor, and factor B.";						
RT	J. Biol. Chem. 263:12403-12411(1988).						
[2]	SEQUENCE FROM N.A.						
RP	MEDLINE=88190151; PubMed=2833753;						
RX	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;						
RA	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";						
RT	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).						
[3]	SEQUENCE FROM N.A.						
RP	MEDLINE=88257215; PubMed=2454931;						
RX	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;						
RA	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";						
RT	J. Cell Biol. 106:2153-2158(1988).						
[4]	SEQUENCE FROM N.A.						
RP	MEDLINE=93123748; PubMed=8419480;						
RX	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;						
RA	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";						
RT	J. Immunol. 150:480-490(1993).						
[5]	SEQUENCE OF 9-1153 FROM N.A.						
RP	MEDLINE=89098893; PubMed=2563162;						
RX	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;						
RA	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";						
RT	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).						
[6]							

SUMMARIES		Description	
Result No.	Score	Query Match %	ID
1	5852.5	99.6	ITAM_HUMAN
2	4476	76.2	ITAM_MOUSE
3	3473	59.1	ITAX_HUMAN
4	3417	58.2	ITAD_HUMAN
5	1552.5	26.4	ITAL_HUMAN
6	1533.5	26.1	ITAM_MOUSE
7	1157.5	19.7	ITAE_MOUSE
8	1149	19.6	ITAE_HUMAN
9	1102.5	18.8	ITAI_HUMAN
10	1093.5	18.6	ITAH_HUMAN
11	1084	18.5	ITAZ_BOVIN
12	1072	18.2	ITAZ_MOUSE
13	1069	18.2	ITAZ_HUMAN
14	1065.5	18.1	ITAG_HUMAN
15	1060	18.0	ITAI_RAT
16	663	11.3	ITAA_MOUSE
17	633	10.8	ITAA_HUMAN
18	630	10.7	ITAG_HUMAN
19	593.5	10.1	ITAA_XENLA
20	571.5	9.7	ITAX_CRISP
21	567.5	9.7	ITAX_MOUSE
22	555.5	9.5	ITAS_MOUSE
23	546.5	9.3	ITAS_XENLA
24	543.5	9.3	ITAV_CHICK
25	538.5	9.2	ITAE_HUMAN
26	535	9.1	ITAV_MOUSE
27	532	9.1	ITAH_HUMAN
28	531.5	9.0	ITAG_CHICK
29	530.5	9.0	ITAE_CHICK
30	526	9.0	ITAV_HUMAN
31	517	8.8	ITAH_CHICK
32	512.5	8.7	ITAE_MOUSE
33	498	8.5	ITAT_MOUSE

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
conservation across species and homology to platelet IiB/IiIIa.";
RL Biochim. Biophys. Acta 874:369-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=96362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
Mutchler V.T., Tonich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=96226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3b. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RWHUB.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3_
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT

Query Match		99.6%; Score 5852.5; DB 1; Length 1152;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 1132; Conservative		3; Mismatches	1; Gaps 1;
Qy	1	FNLDTENAMTFOENARGCQSVQVQGGSRVVVGAPOEIVAAHQSGSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFOENARGCQSVQVQGGSRVVVGAPOEIVAAHQSGSLYQCDYSTGSCPEI	76
Qy	61	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQQPQK	120
Db	77	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVKGLCFPLGNSLRQQPQK	136
Qy	131	FPFALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVWBLKXSLPSLWQYSEEP	180
Db	137	FPFALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVWBLKXSLPSLWQYSEEP	196
Qy	181	RHPTFKFQNNPNRSLVKPTITOLLGRTHATGVRKIVIRBELNITNGARKNAFKILVVI	240
Db	197	RHPTFKFQNNPNRSLVKPTITOLLGRTHATGVRKIVIRBELNITNGARKNAFKILVVI	256
Qy	241	TGGEKFGDPLGYEDVIPADRGVIRYVIGUDAPRSEKSRQELNTIASKPRDRHVFQVN	300
Db	257	TGGEKFGDPLGYEDVIPADRGVIRYVIGUDAPRSEKSRQELNTIASKPRDRHVFQVN	316
Qy	301	NFEALKTIONLRKXTAETEGTQSSSPHEMSQEGPSAATGNGPLLSVWGSYDWAG	360
Db	317	NFEALKTIONLRKXTAETEGTQSSSPHEMSQEGPSAATGNGPLLSVWGSYDWAG	376
Qy	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRVQSLVIGAPYOHIGILVAFR	420
Db	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRVQSLVIGAPYOHIGILVAFR	436
Qy	421	QNTGWESNANUKGTQIGAYFGASLCSVDVDSNGSDTLVGLGAPHYETRGQVSVCP	480
Db	437	QNTGWESNANUKGTQIGAYFGASLCSVDVDSNGSDTLVGLGAPHYETRGQVSVCP	496
Qy	481	PRGORARWQCDVLYGEQCPQGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
Db	497	PRG-RARWQCDVLYGEQCPQGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	555
Qy	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTWDGLVDTVGAQGHVLLRQ	600
Db	556	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTWDGLVDTVGAQGHVLLRQ	615
Qy	601	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLVQKSTRDLRQEQIQSVVT	660
Db	616	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLVQKSTRDLRQEQIQSVVT	675
Qy	661	YDLALDSGRPHSAVFNETHKSTRTOVLGLTQTCETLKLQPLNCIEDPVSPVLRNLF	720
Db	676	YDLALDSGRPHSAVFNETHKSTRTOVLGLTQTCETLKLQPLNCIEDPVSPVLRNLF	735
Qy	721	SLNGTPLSAGNLRPVLAEDDAQLFTALPFPEKNCNDNIQDDLSITPFSMSLDCLVVG	780
Db	736	SLVGTPLSAGNLRPVLAEDDAQLFTALPFPEKNCNDNIQDDLSITPFSMSLDCLVVG	795
Qy	781	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSTRKYSTLQNSQSRWSRLACESSTEV	840
Db	796	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSTRKYSTLQNSQSRWSRLACESSTEV	855
Qy	841	SGALKSTSCSINHPIPRNSEVTFNITFDVDSKASLGKLLKANTVSENMPRTKTRF	900
Db	856	SGALKSTSCSINHPIPRNSEVTFNITFDVDSKASLGKLLKANTVSENMPRTKTRF	915
Qy	901	QLELPVKYAYVMVTHSGVSTKYLNFNTASNTSRVMQHQVQVSNLQGRSLPISLVLPV	960
Db	916	QLELPVKYAYVMVTHSGVSTKYLNFNTASNTSRVMQHQVQVSNLQGRSLPISLVLPV	975
Qy	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAEKARVNCSTAVCQRIQCDIP	1020
Db	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAEKARVNCSTAVCQRIQCDIP	1035

RESULT 2
ITAM MOUSE STANDARD; PRT; 1153 AA.
ID ITAM MOUSE
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RW [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88312584; PubMed=3044779;
RA Pytela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology
RT with the integrin family and an additional domain related to von
RT Willebrand factor.";
RL EMBO J. 7:11371-1378 (1988).
RN [2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,
RA Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse
RT complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).
RN [3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542 (1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES.
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3. THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN
CC MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED
CC GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M
CC SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN
CC RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS. EVENTS THAT
CC APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP
CC OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE. family.
CC -!- SIMILARITY: Belongs to the integrin alpha chain
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07640; CAA30479.1; -;
 EMBL; ML4293; AAA39484.1; -;
 PIR; S00551; S00551.
 HSP; P11215; IABX.
 MGD; MG1:96607; Itgam
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR002035; VWF_A.
 Pfam; PF01839; FG-GAP; 3.
 Pfam; PF00357; Integrin_A; 1.
 Pfam; PF00092; vwa; 1.
 PRINTS; PR01185; INTEGRINA.
 PRINTS; PR00453; VWFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; VWA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Calcium; Repeat.
 SIGNAL 1 16
 CHAIN 17 1153
 DOMAIN 17 1105
 TRANSMEM 1106 1129
 CYTOPLASMIC (POTENTIAL).
 FG-GAP 1.
 FG-GAP 2.
 VWA.
 FG-GAP 3.
 FG-GAP 4.
 FG-GAP 5.
 FG-GAP 6.
 FG-GAP 7.
 CA_BIND 465 473
 CA_BIND 529 537
 CA_BIND 592 600
 SITE 1132 1136
 DISULFID 66 73
 DISULFID 105 123
 DISULFID 654 711
 DISULFID 770 776
 DISULFID 999 1023
 DISULFID 1028 1033
 CARBOHYD 58 59
 CARBOHYD 86 86
 CARBOHYD 391 391
 CARBOHYD 696 696
 CARBOHYD 734 734
 CARBOHYD 772 772
 CARBOHYD 801 801
 CARBOHYD 881 881
 CARBOHYD 907 907
 CARBOHYD 941 941
 CARBOHYD 980 980
 CARBOHYD 994 994
 CARBOHYD 1022 1022
 CARBOHYD 1045 1045
 CARBOHYD 1051 1051
 CARBOHYD 1076 1076
 SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;
 76.2%; Score 4476; DB 1; Length 1153;
 Best Local Similarity 74.3%; Pred. No. 2.6e-290;
 Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

1 FNLDTENAMTFQENARGFGQVYVLOGSRVVGAPQEIIVAAORGSLYOCYSTGSCBPI 60
 17 FNLDTEHPMTQENAKFGQVYVLOGSRVVGAPQEIIVAAORGSLYOCYSTGSCBPI 76

QY 61 RLQVPEAVNMNLSGLSLAATTPPQLLAGCGFVHTQTSENTYVKLCFLPQSNLRQPOK 120
 DB 77 PLQVPEAVNMNLSGLSLAVSTVPQQLLAGCGFVHTQNKENTYVNCVLPQSNLRQPOK 136
 QY 121 PPEALRGCPQEDSDIATLIDGSGSIIIPDPERMEKEFVSTWEOQLFKSKTLFSLMQYSEF 180
 DB 137 PPEALRGCPQEDSDIVFLIDGSGSINNIDPQMEKEFVSTWEOQLFKSKTLFSLMQYSEF 196
 QY 181 RIHFTFXEFQNNPNRSLVKEITQLLGRTHATGVKRVIRELLNITNGARKNAFKILVI 240
 DB 197 RIHFTFNDEKXENPSRSHVSPIKQLNGRTKTASGIRKRVRELPHITNGARENAKILVI 256
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAPRSEKSKOELATITASKPRDHVPOVN 300
 DB 257 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAPRSEKSKOELATITASKPRDHVPOVN 316
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQSFSAITENGPLLSTVGSYDWAG 360
 DB 317 NFEALNTIQNLQEKIFAIEGTQTGSSSPHEHMSQSFSAITENGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTPIIMTRVDSMDADAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR 420
 DB 377 GAFLYTSKDKVTFINTTRVDSMDADAYLGAAAIILNRVQSLVIGAPRYOHIGLVAMFR 436
 QY 421 QNTGWSNANVKTQIGAVFGASLCSVDVDSNGSTDLVIGAPRYOETRGGOVSVCPL 480
 DB 437 ENFGTWEPHTSIKGSQISGYSFGASLCSVDVDSNGSTDLVIGAPRYOETRGGOVSVCPL 496
 QY 481 PRGQARWQCDVLYGEGOGQWGFAGALTVLGVNKGKLTDAVIGAPGEDNRGAVLP 540
 DB 497 PRG-RARWQCEALLHGDQHPWGRFGAALTVLGVNKGKLTDAVIGAPGEDNRGAVLP 555
 QY 541 HGTSGSGISPSHSORIASGLSPRLQVFGQSLGSGQDLTMDGLVDLTVGAOCHVLLRSQ 600
 DB 556 YGASTIASLASHSHRIIGAHFSPGLQVFGQSLGSGQDLTMDGLVDLTVGAOCHVLLRSQ 615
 QY 601 PVLRYKAIMFNPREFVARNFECDQVQVKGKAGEVRVCLHVQKSTRDLREGIOQSIVVT 660
 DB 616 PVLRYKAIMFNPREFVARNFECDQVQVKGKAGEVRVCLHVQKSTRDLREGIOQSIVVT 675
 QY 661 YDLALDGRPHSRVAFENETKSTRTOTVGLTCTETKLQLPNCIEDPSPVLRNLF 720
 DB 676 YDLALDGRPHSRVAFENETKSTRTOTVGLTCTETKLQLPNCIEDPSPVLRNLF 735
 QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFPEKCGNDNICODD:SIITFSMISDCLVVG 780
 DB 736 TLVGEPLRSFGLNRPVLAEDAQRLLFTALFPFPEKCGNDNICODD:SIITFSMISDCLVVG 795
 QY 781 GPRFPNVTVRNDEGDSYRTQVTFPPFLDLSYKRVSTLQNRSCRSWL-ACBSASSTE 839
 DB 796 GPQDFNMSVTLRNDGDSYGTQVTVYPSGLSYKRDASQNPFLTKPWFVKPAESSSSSE 855
 QY 840 VSGALKSTSCSINHPPIPPENSEVTNITFDVDSKASLGNKLLK:NVTSNNMPTNKTE 899
 DB 856 GHGALKSTTWNINHPPIPPANSEVTNITFDVDSHRSFGNKLKLLK:IVASENNMSTRKTK 915
 QY 900 FQLELPVYAVYVWVTSVGVTKYLNFTASENTSRVMOHQVQVSNLQGRSLPISLVFLVP 959
 DB 916 FQLELPVYAVYVWVTSVGVTKYLNFTASENTSRVMOHQVQVSNLQGRSLPISLVFLVP 975
 QY 960 VRLNQTVLWDRPQVTFSENLSTTCHTKERLPSHSDFLAELRKAPVWNCISVACORICDI 1019
 DB 976 VQINNVTVWDRPQVTFSENLSTTCHTKERLPSHSDFLAELRKAPVWNCISVACORICDI 1035
 QY 1020 PFGIQEESFNATLKGNTSFDWYIKTSHNELLIIVSTAEILFNDSVFTLLPGQAGFVRSOTE 1079
 DB 1036 PSFNQKIEFNVTLKGNTSFDWYIKTSHNELLIIVSTAEILFNDSVFTLLPGQAGFVRSOTE 1095
 QY 1080 TKVPEFVNPFLVGVSGVGLLALLIITAAVYKGFPEKQYKMMSEGGPPGAPQ 1137
 DB 1096 TKVPEFVNPFLVGVSGVGLLALLIITAGLYKGFPEKQYKMMSEGGPPGAPQ 1153


```

RESULT 3
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Ileu M5).
GN ITGAX OR CD11C
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8816645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95."
RL EMBO J. 6:4023-4028 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751 (1990).
RN [3]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins."
RL J. Immunol. 138:2381-2383 (1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdic.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81695; AA55180.1; -
CC EMBL; Y00023; CA68283.1; -
CC EMBL; M29165; -; NOT ANNOTATED CDS.
CC EMBL; M29487; AA51620.1; ALT SEQ.
CC EMBL; M29482; AA51620.1; JOINED.
CC EMBL; M29483; AA51620.1; JOINED.
CC EMBL; M29484; AA51620.1; JOINED.
CC EMBL; M29485; AA51620.1; JOINED.

```

```

DR EMBL; M29486; AA51620.1; JOINED.
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR020335; VWFA.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00032; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163 INTEGRIN ALPHA-X.
FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128 POTENTIAL.
FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 165 351 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 POTENTIAL.
FT CA_BIND 530 538 POTENTIAL.
FT CA_BIND 593 601 POTENTIAL.
FT SITE 1131 1135 GPFKR MOTIF.
FT DISULFID 69 76 BY SIMILARITY.
FT DISULFID 108 126 BY SIMILARITY.
FT DISULFID 655 712 BY SIMILARITY.
FT DISULFID 771 777 BY SIMILARITY.
FT DISULFID 848 863 BY SIMILARITY.
FT DISULFID 998 1022 BY SIMILARITY.
FT DISULFID 1027 1032 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490 G -> A (IN REF. 2).
FT CONFLICT 756 756 L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3P62A473 CRC64;

```

Query Match 59.1%; Score 3473; DB 1; Length 1163;

Best Local Similarity 61.3%; Pred. No. 2.1e-223;

Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

Qy 1 FNLDTEAMTFQENARFGSGSVVOLQSGSVVVGAPQPIVANDQCSLYQCDYSTGSCPEI 60

Db 20 FNLDTEBELTAPRVDSAGFSGSVVQYANSVWVGAPQKITAANOTGGYQCGYSTGACEPI 79

Qy 61 RLQVPEAVNMISGLSLAATTPPQLLACGPTVHOTCSENTYVKGCLFGLFSGNLRQOPQK 120

Db 80 GLQVFPPEAVNMISGLSLASTSPQLLACGPTVHHCGRNWLGLCFLIGPT--QLTOR 137

Qy 121 FPEALRGCPQEDSDIAFLDQSGSIIPHDFRMKEFVSTVMEQLKSKTLPSLIQVSEEP 180

Df 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMVFRVAVISQFQRPQSTQSLMQFSNKP 197
Df 181 RHFTFKERONNPNRPSVKPTOLLGRTHRTATGVRKVIPELLNITNGARKNAKILLIVI 240
Df 198 QHFTFEPRRTNPLSLASVHQLGFTTATLQNVVHRFLFHASYGARRDATKILLIVI 257
Df 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVGFDAFRSEKSRQELNTIASKPRDRHVQVN 300
Df 258 TDGKKGESLDVKDVPIMADAAGIIRYAGVGLAFQNRNSWKELDIASKPSQEHFKVE 317
Df 301 NFEALKTIQNLREKIPATEGTQTCSSSEFHEMSEQESFAITSSGILLSTSSVGYDNAG 360
Df 318 DEDALQIQNLQKIKFALIEGTETSSSEFHEMSEQESFAITSSGILLSTSSVGYDNAG 377
Df 361 GYFLVTSKESFTFINNTRVDSMDMAYLGYAAAIILRNVRQSLVGLGAPRYQHIGLVAMFR 420
Df 378 GAFLYPPNMSPTFINMSQENVDNRDSYLCYSTELALWKGVSQSLVGLGAPRYQHTGRAVFT 437
Df 421 QNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVILICAPHYETRCGGVSVCP 480
Df 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTDLVILICAPHYETRCGGVSVCP 497
Df 481 PRGQARMOCDAVLGEGQCPWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Df 498 PRGWR-RWKCDAVLGEGQCPWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Df 541 HGTSGGISPSHSQRIAGSKLRLQYFCQSLGGQDLNMGDLNLTGCAQGVHLLRSQ 600
Df 557 HGVLGPSISPSHSQRIAGSKLRLQYFCQSLGGQDLNMGDLNLTGCAQGVHLLRSQ 616
Df 601 PVLVRKAIJNEPREVARNVFECNDVVKGEAGEVRVCLHVOKSTRDLRGLQSVVT 660
Df 617 PVLWGVSVHQFTPAETPRGAFECREUVSEQTLVQSNICLYIDKSKNLLGSDLSQSVT 676
Df 661 YDLALDSGRPHRAVNETKSTRQTVLGTQTCETKLPNCIEDPVPVILRLNF 720
Df 677 LDALDPGRLSPRATFOETKNSRLSRVVLGLKAHCENFNLLPSCEDSVTPITRLNF 736
Df 721 SVNGTSLARGLNRPVLAEDQRLTALPPEKNCNDMLCDDLSITSEKSLDCLVVG 780
Df 737 TLVGKPELLAFRLNRPMLAALQRYFTASUPPEKNCAGDHIQDNUGISPSFPGLSKLLVG 796
Df 781 GPREFNVTVTRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRMSRLACESASTEV 840
Df 797 SNLELNAEVMWVNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRMSRLACESASTEV 854
Df 841 SGALKTSQCSINPIIPENSEVTFTNTPDVSQASLGNKLLKLVNSENKPRINKTPE 900
Df 855 SQGTWSTSCRIHLEIFRGQAQITFLATFDVSPKAVLGDRLLTANVSSENNTPTSKTTF 914
Df 901 QLELPVKYAVYVTVSHGVSTKYLNFTAS-ENTSRVWQHOYQVSNLQORSLPISLVFLVP 959
Df 915 QLELPVKYAVYVTVSHGVSTKYLNFTAS-ENTSRVWQHOYQVSNLQORSLPISLVFLVP 974
Df 960 VRLNQTVMDRPOVTSSENLSCTHKEKRLPSHDSFLAELKAPVWCVCSINVCORICDI 1019
Df 975 VELNQAVMDVESHVPQNPDLSCSEKSIAPASDFLAHQNPVLDSCSIAGCLRFRCDV 1034
Df 1020 PFGIOEFNATLKNLSRDVYKTSNHLIYSTABILLFNDVSFTLLPQOQAFVRSQTE 1079
Df 1035 PFSVQELDPTLKNLSFGWVQILQKVVSVSVAEITFTSVYSQLPQOQAFVRSQTE 1094
Df 1080 TKVEPPEVPPNPLIVGSSVGGILLALLALITAAALYKLGFFKRYQKMMSE 1128
Df 1095 TULEKVKVHNPTLIVGSSVGGILLALLALITAAALYKLGFFKRYQKMMSE 1143

RESULT 4
ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
ITGAD.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
MEDLINE=961111956; PubMed=8777714;
Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
Stanton D.E., Gallatin W.M.;
"A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
3";
Immunity 3:683-690(1995).
[2]
RP SEQUENCE OF 1-235 FROM N.A.
MEDLINE=20187620; PubMed=10722744;
Noti J.D., Johnson A.K., Dillon J.D.;
"Structural and functional characterization of the leukocyte integrin
gene CD11d. Essential role of S1 and S3";
J. Biol. Chem. 275:8959-8969(2000).
[3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
MEDLINE=96257236; PubMed=8666289;
Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
"Cloning and chromosomal localization of a novel gene encoding a human
beta 2-integrin alpha subunit";
Gene 171:291-294(1996).
[4]
RP INTERACTION WITH VCAM1.
MEDLINE=95059842; PubMed=9841932;
Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
Hoffman P.A., Stanton D.E., Bochner B.S.;
"alpha2 integrin is expressed on human eosinophils and functions
as an alternative ligand for vascular cell adhesion molecule 1
(VCAM-1)";
J. Exp. Med. 188:2187-2191(1998).
[5]
RP INTERACTION WITH VCAM1.
MEDLINE=95370002; PubMed=10438935;
Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
Grayson M.H., Bochner B.S., Gallatin W.M., Stanton D.E.;
"The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
binding interface between I domain and VCAM-1";
J. Immunol. 163:1984-1990(1999).
-!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES
FROM THE BLOOD.
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
ASSOCIATES WITH BETA-2.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FROM CELLS WITHIN
ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
-!- SIMILARITY: Belongs to the integrin alpha chain family.
-!- SIMILARITY: Contains 1 VWFA domain.
-!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@sib-sib.ch.

EMBL; U37028; AAB38547.1; -
DR EMBL; U40274; AAB60634.1; -
DR EMBL; U40275; AAB60635.1; -
DR EMBL; U40276; AAB60636.1; -
DR EMBL; U40277; AAB60637.1; -
DR EMBL; U40278; AAB60638.1; -
DR EMBL; U40279; AAB60639.1; -
DR EMBL; U40280; AAB60640.1; -
DR HSSP; P11215; IABX.
DR Genew; HGNC:6146; ITCAD.
DR MIM; 602453; -
DR GO; GO:0008305; C: integrin complex; TAS.
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.
DR GO; GO:0016337; P: cell-cell adhesion; NAS.
DR GO; GO:0007160; P: cell-matrix adhesion; NAS.
DR GO; GO:0006955; P: immune response; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF03357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02344; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium;
KW Magnesium.
FT SIGNAL 1 17
FT CHAIN 18 1162
FT DOMAIN 18 1100
FT TRANSMEM 1101 1124
FT DOMAIN 1125 1162
FT REPEAT 32 95
FT REPEAT ? ?
FT DOMAIN 150 332
FT REPEAT 350 400
FT REPEAT 401 452
FT REPEAT 454 516
FT REPEAT 518 576
FT REPEAT 581 633
FT CA_BIND 465 473
FT CA_BIND 530 538
FT CA_BIND 593 601
FT SITE 1127 1131
FT DISULFID 67 74
FT DISULFID 106 124
FT DISULFID 655 710
FT DISULFID 769 775
FT DISULFID 846 861
FT DISULFID 994 1018
FT DISULFID 1023 1028
FT CARBOHYD 59 59
FT CARBOHYD 87 87
FT CARBOHYD 99 99
FT CARBOHYD 391 391
FT CARBOHYD 691 691
FT CARBOHYD 733 733
FT CARBOHYD 873 873
FT CARBOHYD 957 957
FT CARBOHYD 1046 1046
FT CONFLICT 500 500
FT CONFLICT 515 518
FT CONFLICT 825 825
FT CONFLICT 984 984
FT SEQUENCE 1162 AA; F296AIA35455D77D CRC64;
Query Match 58.2%; Score 3417; DB 1; Length 1162;
Best Local Similarity 59.8%; Pred. No. 1.1e-219;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;


```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT VARSPLIC 954 PWGVSSEPRDNIRAGPCR (in isoform 2).
/FTid=VSP 002738.
R -> W (IN REF. 1 AND 2).
Y -> I (IN REF. 2).

214 214
660 660
155 162
164 164
165 166
169 185
186 188
191 198
202 206
208 213
217 221
222 223
229 229
233 243
244 244
247 249
250 250
253 254
256 263
274 276
277 278
280 286
288 290
293 297
298 298
300 302
307 310
311 314
318 318
319 328
329 330
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EP286FC0 CRC64;

Query Match
Best Local Similarity 26.4%; Score 1552.5; DB 1; Length 1170;
Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;

1 FNLDTENAMTFO--ENARGFGSVVQLOQSGRVVVGAPQEIIVANORGSLYQCDYSTGSGCE 58
26 YNLDVRGARSFSPRAGRHRFGYRLQV--GNGVIVGAPGE--GNSGTGSLYQCQSGTGCHCL 81
59 PRLQVPVEAVNMSGLSLAANTSPQLAACGPTVHCTSENTYKGLCFPLGSLNR--- 115
82 PVTLR--GSNTYKYLQNTLADPTDGSIIACDPGLSRICDQNTYLSGLCYLFRQLQGPW 140
116 -QOQKFPBALRCPOEDSDIAFLIDSGSIIPHDFRMRKMFVSTVMBQKSKTLFSLM 174
141 LQGRPGFCIKG---NVDLVPLFGSSMSLQDFEFQKILDFMKDWKLSNTSYQFAAV 196
175 QVSEEPRIHPTKPEFQNNPNRSLKVPITQLLQRTHTATGVRKYVRELLNITNGARKNAF 234
197 QESTSYKTEFDPSDYVYKRPDPDALLKHVHMLLTNTFGAINVATEVREBELGAPDAT 256
235 KILVITDGEKFDPLGYEDVTPEDREGVIRVIGVGDAFRASEKSRQELNTIASPPPD 294
257 KVLIIITDGE--ATDGSNDIAKD-----IIRVIIGKHFQKESQETLHKEFSKASE 309
295 HVFQVNNFRALKTIQNLQREKIPIAIEGTQGTGSSSPFHEHMSQEGFSAALITNGPILSTVG 354

```

RESULT 6

ITAL MOUSE

ID ITAL MOUSE STANDARD; PRT; 1163 AA.

AC P24063;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1

DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)

DE (CD11a).

GN ITGAL OR LFA-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Db 310 FVKILDTPEKLDLPTELKOKIIVIEGTSKQDLTSENMELSSSGIADLSRGHVVAVG 369
QY 355 SYDNAGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAAA-IILENRVOSLVLGAPRYOH 412
Db 370 AKDNAGGFDLKDADQDDTFIGNEPITPVRAGYLGITVTWLPSPKXTSLASGARYOH 429
QY 413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSITDLVIGAPHYEQT 470
Db 430 MGRVLLFQBPQGGCHMSQVQTHGTQIGSYFGELGVDVDDQGETELLIGAPLFYGEQ 489
QY 471 RGOQSVCLPLPGORARWQCDAV--LYGSGOGPWGFRGAALTVLGVDNGDKLTDVAIGAP 528
Db 490 RGRGRVIY-----QRKQLGFEEVSELQGDPGYPLGRFGEAITHALTIDNGDLVDVAGAP 544
QY 529 GEEDNEGAVYLFEGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLMDGLVDLTV 588
Db 545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEQTQVSGIQWFGRSIHGVKDLGEGDLADVAV 601
QY 589 GAQCHVILLRSQPVLRVKALMEINPREVARNVTECNDQVY-KGKEAGEVTVCLHVOKSTR 647
Db 602 GAESOMIVLSRSEFVDMVTILMSFPAEIPVHEVECSYSTSNKMKEGVNIITICQI-KSLY 660
QY 648 DRLREGIQSVVTVYDLALDSGRPHSAVENETKNSRTERQTQVLGLTOTCETLKLQFNCI 707
Db 661 PQP-QGRLVANLYTLQDGHRTRRGLPGGRHLEARNIAVT-TSMSCDTDFSHFPCV 718
QY 708 EDPVSPIVLNFSL---VQTPLS--AFGN-----LRPVLAEADQRLFTALFPKRCGN 757
Db 719 QDLTSPINVSINFSLMBEECTPRDQRAQKDIIPILRPSLSHSETWEL-----PEKNCGE 773
QY 758 DNICQDDLSITPSFMSLDCLVVGPGREFNVTVTVRNDGDSYKTCVTFPPPLDLISYRKVS 817
Db 774 DKCEANLRVSFSPARSARALRLTAFASLSVELSLSNLEEDAYVWVCLDLHFPPLGSLPKVE 833
QY 818 TLQNRQSRWRLACES--ASSTEVSGALKSTGCSIMHPIPPENSEVTFNITPVDVSKAS 875
Db 834 ML---KPHSQIPVSCBELPEESRLSRAL--SCNVSSPIFKAGHSVALQWMTLVNSS 887
QY 876 LGNKLILLKANTYSENN---MPTNKTQFQLELPVKYAVVMVMTSHGVSTKYLNFIASEN 931
Db 888 WGSDELHANVTCTNEDSDLEDNSATFI---IPILPINILIQDQEDSTLYVSFTKGP 944
QY 932 TSRVMOHQYQV---SNLQQRSLP-ISLVFLVFLVNLQTVIWDPRQVTFPSENLSTCHTK- 986
Db 945 KIQVKHNYQVRIQPSIHENIPTLEAVGVFPQPSGPIHQWSVQMEPPV--PCHYED 1002
QY 987 -ERLPSSHSD--FLAELRKAPVWNCISIAVCQRIQCDIPFGIQEENFATLKNLSFDMYIK 1043
Db 1003 LERLPDAAEPCLPGALFPCPV-----FRQELLVQVIGTLELVGEIE 1044
QY 1044 TSHNHLIVSTAELIFNDVSFTLLPGQAFVRSQTEKVEPPEVFNPLFLIVGSSVGLL 1103
Db 1045 AS-SMPSLCSLSISFNSSKHFLYGSNASL-AQVWVKVDVVYKQMLYLVLSIGIGLL 1102
QY 1104 ILALITAILYKGLPFKQYKDMMSSEG-GPPGAEP 1136
Db 1103 LLLLIIFVLYKVGPFKRLKKEKMEAGRVPNGIP 1136

```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91268576; PubMed=2051027;
 RA Kaufmann Y., Tseng E., Springer T.A.;
 RT "Cloning of the murine lymphocyte function-associated molecule-1
 alpha-subunit and its expression in COS cells.";
 RL J. Immunol. 147:369-374(1991).
 RN [2]
 RP SEQUENCE OF 24-42.
 RX MEDLINE=85188276; PubMed=3887182;
 RA Springer T.A., Teplov D.B., Dreyer W.J.;
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 RL Nature 314:540-542(1985).
 CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
 CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
 CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
 CC LEUKOCYTES RECRUITMENT.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
 CC -!- WITH 1-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60778; AAA39426.1; -.
 DR PIR; I56126; I56126.
 DR HSSP; P20701; 1LFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; WVF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; WVPADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium;
 KW Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 1163 INTEGRIN ALPHA-L.
 FT DOMAIN 24 1084 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1085 1108 POTENTIAL.
 FT DOMAIN 1109 1163 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 39 98 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 148 334 VWFA.
 FT REPEAT ? ? FG-GAP 3.
 FT REPEAT 399 454 FG-GAP 4.
 FT REPEAT 455 514 FG-GAP 5.
 FT REPEAT 516 573 FG-GAP 6.
 FT REPEAT 576 628 FG-GAP 7.
 FT CA_BIND 466 POTENTIAL.

FT	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GPGR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	108	126	BY SIMILARITY.
FT	DISULFID	147	199	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	1163	1163	AA; 128343 MW; A7A3078489E8232F CRC64;
QY	Query Match	26.1%;	Score 1533.5;	DB 1; Length 1163;
Db	Best Local Similarity	34.0%;	Pred. No. 4.7e-94;	
QY	Matches	398;	Conservative 217;	Mismatches 458; Indels 99; Gaps 36;
QY	1	PNLTENAMTFQENVA-RGFGQSVVLOGSRVNVGAPQEIIVAAQSGISLYOCDYSTGSCPE	59	
Db	24	YNLRTPTQFLAAGHFGYQVLQIEDG-VVVGAPGE---GDNLTGLMHCRTSEFCQP	79	
QY	60	IRLQVPVEAVNMSLGLSLAATTSPPLLACQPTVHQCSTNTYVYKGLCFLFGSLRQOQP	119	
Db	80	VSLH-GSNHTSKYLGWLTATDAKGSLLACDPLSRCTDQNTYLSGLCYLPQSLQEPML	138	
QY	120	KFPALRGCPQEDSDIAFLIDGSGIIPHDPRMKPFVSTVMEQLKSKTLFSLMQVSEE	179	
Db	139	QNPAYQECMKGRVDLVFLDFGQSQSLDRKDFEKLFEPMKVMRKLSTNTSVQFAAVQST	198	
QY	180	FRIHTEKEP-ONNPERSIVKPIITOLLGTHATGVRKVIKRELLNITNGARKNAFKILI	238	
Db	199	CRTEFTFDYVKQKNEVDVLGVSQPMFLINTFRANVYVAHVKEESGARPDATKVLV	258	
QY	239	VITDGEKFGDPLGYEDVIPADREG-----VIRYVIGVGDAPRSEKSGOELNTIASKP	291	
Db	259	IITDG-----EASDKGNISAAHDITRYIGIGKHFVSVQKQKTLIFASEP	304	
QY	292	PRDHVQNNFEALKTIQNLREKIPALEGTQGTSSSSPEHEMSCQEGFSAITNGPLLS	351	
Db	305	VEBFVKILDTFEKULDLTDLQRRYIALEGNTRODLTSFNMELSSGISADLSKGHAVVG	364	
QY	352	TVGSYDAGGVF-LYTSKEKSTFINMTRVSDMNDAYLGYAAA-IILNRVQSIVLAGAPR	409	
Db	365	AVGAKDWAGGFLDLREDLQCATFVGQEPILTSDVRGGYLGTVAWWTSSRSRPLLAAGAPR	424	
QY	410	YQHIGLYAMER--QNTGMWSSNNVKGTOIGAYFGASISVDVDSNGSTDLVLIGAHY	467	
Db	425	YQHVQVLLFOAPAGGRWNTQKIEGTQIGSYFGELSCVDDLCDEAEALLIGAPLFP	484	
QY	468	BQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALFVLGVNWDKLTDAIGA	527	
Db	485	GEQGGGRVFTY---QRRQSLFEMVSELQGDGPVPLGRFGAAITALTIDNGDRLTDVAVGA	541	
QY	528	PGEEDNRGAVYLFGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLDLT	587	
Db	542	PLEE--QGAVYIFNGKFG-GLSPQPSQRIQGAQVFPGRIMFGRSIHGVDLGGDRLLADV	598	
QY	588	VGAQGHVLLLRSPVLKVKALMEFNPREVARNVECDNQVVKGEAG-EVRVCLFVQKST	646	

FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MM; 88331C115DCCCFD CRC64;
Query Match 19.7%; Score 1157.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 5.7e-69;
Matches 358; Conservative 211; Mismatches 459; Indels 205; Gaps 43;
QY 1 FNLDTENA--MTFQENARGFQSVVQLOGRVAVVGAPOEIVAAQNGRS-----LYQCDY 52
DB 20 FNMVDVWAVTALQAGAPVAVLSLLHDPN-----NOTCLLVARRSNRNTAALYRCAL 74
QY 53 STGSCPEIRLQVPEAVANMGLSLAATT--SPQALLAC-GPTVHQTCESTYVKGICFL 109
DB 75 SI-SPDEIACO-PVEHICMPXGRGVGTLVGNHNGVLVCIQVQARKFRLNSELTGACSL 132
QY 110 FGSNLRQOPQPEALRG-----C-----POE 131
DB 133 LTPNLDLQAAVYFSLGFLDGPAGVDSGDYCRSKGSGTGEEKSARRRRTVBEEDREED 192
QY 132 DSDIAFLIDGSGSIIPHPFRMBKPVSTVWEOQ--KSKTLFSLMOYSEEPRIHFTKEF 189
DB 193 GTEIAVLIDGSGSIIPHPFRMBKPVSTVWEOQ--KSKTLFSLMOYSEEPRIHFTKEF 252
QY 190 QNNPNRSLVXPIITOLLGRTHATGVRKVIKRELLMINTGARKNAFKLIVITGKEFGDP 249
DB 253 RDINASLAKVQSIQVQKEVKTASAMQVLDNIFIPSRGSRKALKVWVLTGDIQFDP 312
QY 250 LGEYDVPEADREGVIRVIVGVDAFRSEKSEQLNTIASPPRDHVFQVNFALKTIQ 309
DB 313 LNLITVINSPMQGVVRFAIGVDFKNNRTRELKLIASPKEMHTPKVINYALDGLL 372
QY 310 NQLEKIFAIETQGTSSSEHMSQSFSAITSNGP--LLSTVSGYDWAAGVFLY--TS 367
DB 373 SKLQQRIVMEGT---VGDALQYQLAQTGFSAQILDKGVLLGTGVAFWNSGGLLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMDAYLGVAAAILNRRVQSLVLAQPRYOHGLVAMERQVGM 425
DB 430 NRGCFNLCTAKEDSRTOYSYLGSVLAHLKHAHGISYVAGAPRHURGAFTELKEDR- 488
QY 426 WESNA---NVRGTQICAFYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGQVQVCPLEP 482
DB 489 -EEDAFVRIEGBQWGSYFSGVLCVFDIDMDGTTDFLLVAAPFYHIRGEGRVYVQVE 547
QY 483 GORARWQCDVLYGQGFQGRFGAALTVLGVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLSAHTLSGHFGTNSRFGFAMAAGVDINQDKFTDVAIGAPLEGFGAGDGASYS 606
QY 537 VYLFHGTSGSISPSHSGRIAGSKLSPLQYFGSLGSGQDLTMDGLVDLTWGAQGHVLL 596
DB 607 VYINGHSG-GLYDSFSQIIRASSVAGSLYFGHVSGLDFNGDLADIIVGSRDSAVV 665
QY 597 LRSQPLVAKVIMFNPREVARNVPECDQVVKGEAGEVRVCLHVQKS-----TDRRLREG 653
DB 666 LRSRPVLDLTVSMFTF-----DALPMVPIGRH--DVNLCEPVDSSVVASBFLREM 715
QY 654 QIQSVTVYDLDLDSGRPHSRVFNENKSTRQTOVLGJLTC-----696
DB 716 FLNFTVDVU-----TKORQLQCEDDSGGSGCLKXNGSGSLCEHFWLI 760
QY 697 ETLKQLPNCIEDPVSPVLRNLFSLVGTPIPSAFGNLR-----PVLAEADAQLFTALP--P 750
DB 761 STEEL-----CEDDCFSNITIKVYE-----FOTSGGRDYENFTL--DHYKEPSAIFQLP 809
QY 751 FEKNGCNDNIQODLSLITFSMLDCLVAVGGPREFNTVTVNRDGEDSVYRTOVTFEPLD 810
DB 810 YEKQCNKVFCAIQLITN--ISQOELVGVGVTKVTMMISUNSGEDSMTNMAALNPRN 868
QY 811 LGRYKVTILQNRQSRMLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITPDV 870
DB 869 LQFKKI-----QKPVSPDVQCDPKPV---ASVLWACKIGHPII--KRSSVNVSVTWQL 918

QY 871 DSKASLGNKLLKANVTSENNMPTNKTFQLELPVKYAVVMVTSHGVSSTKYLNFITASE 930
DB 919 EBSVFNPRADIVTIVTSNKSXLAEETR---SLOFHFHAFIAVLSR--PSVMTWN--TSQ 971
QY 931 NTSRVNQHQYQVNSNLGORSPLISLVLVPLVRLNQTVIMDRPQVTFSENLSST-----CHT 985
DB 972 SPDSHKFEFFNVHGENLFGAVFQIQICVPIKLQDP-----QIVRVKXNLTKQDHTCTQ 1025
QY 986 KERLPSSHDSPLAELRKAPVNVCSIAVCRIQCIPFFGIGQEEFNATLKGMLSDWDYIKTS 1045
DB 1026 SQSPACGSDPVQHVKSHSVCAI-----TSKNENVTVAAEISVG 1065
QY 1046 HNHLIVSTA-----EILFNDVSFTLLPQGGAFVRSQETKVEFP-----EVNPLPLIV 1095
DB 1066 HTKQLLRDVSSELPIGLSIFSNKSLVEGLNAE-----NHRKTIVTIFLKEERTSLPLII 1119
QY 1096 GSSVGGILLALITALYKLGKFKKROVKDMKSE 1128
DB 1120 GSSIGGLLVLLVIAILKCGFFKRYQQLNLZ 1152
RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1
DE antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
[1]
RN _SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
EX MEDLINE=94164962; PubMed=8113947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
RA Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E
RT subunit. Unusual structure and restricted RNA distribution.";
RT J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN _SEQUENCE OF 53-1179 FROM N.A.
RP TISSUE=Petal kidney;
RC MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,
RA Stoteleruk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
RA Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a
RT novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RN _MUTAGENESIS OF ASP-109 AND PHE-316.
RP MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
RA Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
CC EPITHELIAL CELLS.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L25851; AAB59359.2; .
CC EMBL; AF158787; AAF43107.1; .
CC PIR; A53213; A53213.
CC HSSP; F1215; IASX.
CC GENE; HGNC:6147; ITGAE.
CC MIM; 604682; .
CC GO; GO:0008305; C: integrin complex; TAS.
CC GO; GO:0004895; F: cell adhesion receptor activity; TAS.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWFA.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Ict_alpha; 3.
CC SMART; SM00327; VWFA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS50334; VWFA; 1.
CC InterPro: Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
CC SIGNAL 1 18
CC CHAIN 19 1179 INTEGRIN ALPHA-E.
CC CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.
CC CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
CC DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1125 1147 POTENTIAL.
CC DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 181 198 GLU-RICH (ACIDIC).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 200 391 VWFA.
CC REPEAT 401 456 FG-GAP 3.
CC REPEAT 457 506 FG-GAP 4.
CC REPEAT 510 571 FG-GAP 5.
CC REPEAT 573 638 FG-GAP 6.
CC REPEAT 641 693 FG-GAP 7.
CC CA_BIND 522 530 POTENTIAL.
CC CA_BIND 586 594 POTENTIAL.
CC SITE 1150 1154 GFPR MOTIF.
CC DISULFID 70 79 BY SIMILARITY.
CC DISULFID 126 159 BY SIMILARITY.
CC DISULFID 706 762 BY SIMILARITY.
CC DISULFID 823 829 BY SIMILARITY.
CC DISULFID 893 907 BY SIMILARITY.
CC DISULFID 1008 1033 BY SIMILARITY.
CC DISULFID 1041 1057 BY SIMILARITY.
CC CARBOHYD 49 409 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).

FT	CARBOHYD	726	726	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	934	934	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	954	954	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1065	1065	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	360	360	D -> E.
FT	VARIANT	1041	1041	/FTid=VAR_008884.
FT	MUTAGEN	208	208	/FTid=VAR_008885.
FT	MUTAGEN	316	316	D->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	477	477	F->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	482	482	V -> I (IN REF. 3).
FT	CONFLICT	950	950	Q -> R (IN REF. 3).
FT	CONFLICT	1019	1019	R -> W (IN REF. 3).
FT	CONFLICT	1179	1179	A -> V (IN REF. 3).
FT	SEQUENCE	1179	1179	AA; 130088 MW; E558902EDP9D5E1 CRC64;
Qy	45	GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPOLLACGPTVHQ	95	
Db	65	GPLHRCSLVQDEILCHPVEHVPKGRHGVTVVRSHGVLCI-----QVLVRP--HS	117	
Qy	96	TCSENTYVKGCLFLGSLNRQQPQ-----	119	
Db	118	LSSELT---GTCSLGLPDLRPAQANFFDLENLLDPARVDTGDCYSNKEGGEDDVNTA	174	
Qy	120	KPPEALRGCTQED-----SDIAFLDGGSLIPHOFRMKRFVSTVMEQL--	164	
Db	175	RQDALEKEEEDKEEEDBEAEAGTEIAILLDGGSIDPPDFCRANDFSNNWRFYE	234	
Qy	165	KKSKTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGVKVRIBLLN	224	
Db	235	KCFECNALVOYGGVIOQTEFDLRDSQDVMAASLARQNTVGSVTKTASAMQHVLDST	294	
Qy	225	ITNGARKNAKILIVITDCKFKGDPGLGYEYVIPADREGVIRYVIGVDARESEKSRQEL	284	
Db	295	SSGSRKASKWVVLTDGGIFEDFLNLTTVINSPKQGVRFPAIGVGBEFKSAARTAREL	354	
Qy	285	NTIASKPRDHVFOVNNFEALKTIONLRKIPIAEGTGTGSSSSFHEMSQEGFSAIT	344	
Db	355	NLIASDDETHAFKVTNYMALDGLSKLYNIISWEGT---VGDALHYLAQIGFSAQIL	411	
Qy	345	SNGP-LLSTVGSYDWAGGVFLY-TSKEKSTFTNNTRVDSMDNDA----YLGYAAAILRN	398	
Db	412	DERQVLLGAVGAPDWSGALLYDTRSRGRFLNTAAADAAEAQYSLGYAVAVLHK	471	
Qy	399	RVQSLVLCAPYQHIGLVAMFR-QNTGMWESNANV-KGTQIGAVFGASLCSVDVDSNGST	456	
Db	472	CSLSYVACAPQYKHG--AVPELOKREASFLVLEGEQKGSYFGSELCPVDIDMGST	529	
Qy	457	DLVLIGAPHYYEOTRGQVSVCPFRQQRARQCDAVLVYGEQGVWGRFGAALTVLGDVN	516	
Db	530	DFLLVAAPFYHVHGEGRVYVRLSE-QDGSFSLARILSGHPGTNARFGPMAAMGDLS	588	
Qy	517	QDKLTDVAIGAP-----GEEDNR--GAVYLFHTSGSGISPSHSOFIAGSKLSPLRYEQG	570	
Db	589	QDKLTDVAIGAPLEGFGADGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM	647	
Qy	571	SLSGGQDLTMDGLVDLTVGAQGHVLLLRSSQPLVRKATMEFNPVARNVFCNDQVYKG	630	
Db	648	SMAGGPDISGDLADIIVGTIGQAVVFRSRPVRLKVSMAFTPSALP-----IGF	697	
Qy	631	KEAGEVRVCLHVQKSTRDLRBEQIOQSVVTDLALDGRPHRAVFNFTKNSRQTVL	690	
Db	698	NGVNVRLCPFI-SSVTATSSGRLREALNFTLDVGVKQRRRLQCSDVRSCLGCREWS	756	
Qy	691	GLTQTCETLKLQLEN-----CIEDPVSPIVLKRLNLSLVTPLSAFGLNLRPLVLAEDAQRLFT	746	

Db 757 SGSQCEBL-LIMTEGELCEBDFCSNASVKVSYQL-QTPBQQTDPQPILDRTYTFEFAI 814
Qy 747 ALPPEKNCNDNQCDDLSITFSFMSLDCLVNGSPREFNTVTVRVDGEDSYETQVTF 806
Db 815 FOLPYEKACKNKFVAVELQLA-TTVSQEUVGLTKELTUNINELTNGSDSYMTSWALN 873
Qy 807 FPLDLSYRKVSTLQNRQSRWRLACESASSTVSGALKSTSCSINHPIPPENSEVTNFI 866
Db 874 YERNLQ-----LKMQRKPPSPNIQCDPQPV---ASVLIWNCIRGHFVL-KRSSAHVSV 923
Qy 867 TFDVDSKASLGNKLLKANVTSENN-----MPTNKTEFO---LELPVKYAYVWVTSHG 919
Db 924 VQLEENAFNPRTADITVTVNSNERRSLANETHLOFRHGFVAVLSKPSIMYNTQOGL 983
Qy 920 S--TKYLNFTASENTSRVMQHQYQVSNLQSRSLPISVFLVPLVFNQTVIMDRPQVTFSE 977
Db 984 SHKKEFLPHVHGEN--LFGAEYQ-----LQICVPTKLGQLQVAAVKKLRITQ 1028
Qy 978 NLSSTCHTKERLPHSDFLAELRKAPVNVCSINVCRIQCIDIPFGIOBEFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVEMHVSVCVIA-----SDKENVTVAAEIS 1073
Qy 1038 FDWIKTSHNLLTVST-----ABILFNDVSFTLLPGQAFVRSQVETKVEPPEVNPPL 1091
Db 1074 WD-----HSELLKDVTELQILGEISFNKSLYBCLNENH--RTKITVWELKDEKYS 1125
Qy 1092 PLIVGSSVGGLLLLALITAAALYKLGFPKQYKMMSE 1128
Db 1126 PIILGSSVGGLLVILVILFKGCFKRYQQLNLE 1162

RESULT 9
ITAL HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.B.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit.";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1OC5; 17-MAY-00.
DR GENE; HGNC:6134; ITGAL.
DR MIM; 192968; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002035; VWFA_A.
DR Pfam; PF01839; FG-GAP; 3-
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure
PT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
PT TRANSHEM 1114 1136 POTENTIAL.
PT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA BIND 470 478 POTENTIAL.
FT CA BIND 552 560 POTENTIAL.
FT SITE 614 622 POTENTIAL.
FT SITE 1139 1142 GPEPR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 693F3C1AABF52808 CRC64;
Query Match 18.8%; Score 1102.5; DB 1; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.6e-65;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;
Qy 1 FNLDTENAMTFQENARG-FQSVVQL---QGSVVVVGAPQEIIVAANQSGSLYQCDYSTGS 56
Db 1 FNVDKNSMTFSGFVEDMFGYTVQQYENESGKVLIGSLVVGQPKNRTGDVYKCPVGRGE 60
Qy 57 CEP-IRLOVPEA-----VNMSGLSLAATTSPPQLLAGPTVHQCSENTYKGL 106
Db 61 SLPCVKLDLFLVNTSIPNVTEVKNMTFGSL-VTNPNNGGFLACGPLYAYRCGLHYTTGI 119
Qy 107 CFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHPF3RMKEFVSTVWEQLK- 165

Db 120 CSDVSPFTQVNSIAP--VQSCSTQ-LDVIIVLDGNSIYFWD--SVTAFNLDLKRMXI 174
 Qy 166 -KSKTSLMOYSEFRIHFTFKFQNNPRSLVKPITQLLGR-THTATGVKRVIRELL 223
 Db 175 GPKQTQVIGVYGENVTHEFNLYKSYSTEVEVLVAKKIVQGGQRTWALGTDTARKEAF 234
 Qy 224 NITNGARNAPKILIVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAFR-----SE 278
 Db 235 TEARGARGVKWVIVITDGEH-DNHELKKVQDCDENIQRFSAIILGYSYNGNLSTE 293
 Qy 279 KSRQELNTIAPKPRDRHVQNNFEALKTIONQREKIFALEGTQTSSSSPEHEMGOE 338
 Db 294 KFVEEIKSIASEPTKEKFPNVSDELAVTIIVKTLGERIFALEADQSAASFEMKQ 353
 Qy 339 FSAALTNSGPLLSVGSYDAGGVFLYTSKE-----XSTF-INNRVDSMDNDAYLGAA 392
 Db 354 FSAHYSQDWMLGAVGADWNGTVVMQKASQIIPRNTTFNVFSTKXNEPL-ASYLGTV 412
 Qy 393 AILLRNRVOSL-VIGAPYOHIGLVMPQNTGHWESNANKVGTQIGYFGASLCSVDVD 451
 Db 413 NSATASSGDVLYIAGQPRYNHTGVIIYRMDGNKILQTLGSGEQIGSYFSGILTTDID 472
 Qy 452 SNGSTDIVLIGAPHY-----YEQTR-GGQVSVCPPLPRQORARWOCDAVLY 495
 Db 473 KDSMTDILLVGAPMNGTEKEEGKGVVYVALNTRFQYOMLEPIKQCCSRQNSCTT 532
 Qy 496 GEOGQPWG-RFGAALTVLGVNGDKLTDVAIGAPBEDNRGAVLFGHTSGSGISPHSQ 554
 Db 533 ENKEFOGARTGTAFAAVKDLNDGENDIVIGAPLEDHGGAVIYHG-SGKTIRKEVAY 591
 Qy 555 RIASKLSPLRYEQCSLSGGDLTMDGLVDLTGVAQGHVLLLSQPLVRLKALWENPR 614
 Db 592 RIPSQGDGKTLKFFQGSIHGMDLMDGLDVTGIGLGGAALEFWRDVAVVKVNTNFE 651
 Qy 615 EVARNVFCNQVVKVKEAG--EVRVCLHVQ-KSTRDLRREGQIQSVVYTDLALDSGRPH 671
 Db 652 KVIQKQKCH--MEGKETVCINATCFEVLKSKEDTVIYADLQ---YRVTLDLSLRQ 704
 Qy 672 SRAPVNET-----KXSTRQTOVLGLTQTCETLKLQLPNCIEPVPVILVRLNLSLWT 725
 Db 705 SRSPFGTQERKVORNIIVRSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNLTD 755
 Qy 756 PLSAPGNLRPLVLAEDQRLFTALPFKXNCGNDNICQDLSITTSFMSLCLVWGGPRE- 784
 Db 756 PENG-----PVLDDSLPNSVHEYIFPAKDCGKKEKICISLHLVATTEKDLIVRSQNDK 810
 Qy 785 FNVTVVRNDBDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSRWKLACESASSTEVSGAL 844
 Db 811 FNVSLTVKNTKDSAYNTRTIHVHSPNLVFSGTEATQKD-----SCESN----- 853
 Qy 845 KSTSCSINHPIPPENSEVTENITDQVSKASIGN-KLALKANVTSENNMPTNKTTEFOLE 903
 Db 854 HMITCKVGPFLRRGEMVTKILFQNTSYLNMENVTIVLSATSDSEEPPELSDNVNMS 913
 Qy 904 LPVKYAVYVVTSHGVSTKYLNTASENTSRVMQHQYQVSN-----LGQRS----- 950
 Db 914 IPVKYEVGLQFYS-SASEVHISIAANETVPEINSTEIDGNEINFPYLIRKSGSPMDEL 972
 Qy 951 PLSLVP-----LVPRLNQTVIWRDPQVTFSENLSSCTHYE-----RLPS 991
 Db 973 KLSISFPNMTSNGYVLYPTGLSS-----SENANCRPHIFEDPFSINSKQVTT 1021
 Qy 992 HSDFLAELKAPVWVNSIAVCQRIQDIPFFGIQE-----EFNATLK 1033
 Db 1022 STD---HLKRGTLQCNCKPATITCNLTSSDISQVNVSLILKWKTFKSYFSSNLATIR 1078
 Qy 1034 GNLSDFWYIKTHNHLILVSTAEILFNDSVFTLLPQOQAFVRSQVETKVEPEVNPPL 1093
 Db 1079 GEL-----RSENASLVLSGN-----QKRELAIQIKDGLGPRVPL 1114
 Qy 1094 --IVGSSVGLLILALITAAVLKGLFKFQYKDMSE 1128

Db 1115 WVLISAPAGLLMLLILALWKIGFFKPLKKMKX 1151
 RESULT 10
 ID ITAH HUMAN STANDARD; PRT; 1189 AA.
 AC Q9UKS; Q9UKQ1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-11 precursor.
 GN ITGAL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Petal heart, and Osteoblast;
 RC MEDLINE=99417678; PubMed=10486209;
 RX Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
 RA Wang S.-X., Morris C.M., Krissansen G.W.;
 RA "Cloning, sequence analysis, and chromosomal localization of the novel
 RT human integrin alpha11 subunit (ITGAL1).";
 RT Genomics 60:179-187(1999).
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=fetal muscle, and Uterus;
 RC MEDLINE=99395147; PubMed=10464311;
 RX Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
 RA "cDNA Cloning and Chromosomal Localization of Human alpha(11)
 RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated
 RT integrin alpha-chain present in muscle tissues.";
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [3]
 RP SEQUENCE OF 954-1188 FROM N.A.
 RC TISSUE=Fibroblast;
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
 CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
 CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
 CC FETAL MUSCLE CELLS (IN VITRO).
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF109681; AA01258.1; --
 CC EMBL; AF137378; AA05191.2; --
 CC EMBL; AL359064; CAB94392.1; --
 CC HSSP; P17301; 1A0X
 CC Genew; HGNC:6136; ITGAL1.
 CC MIM; 604789; --
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.

DR GO: GO:0005518; F: collagen binding; TAS.
 DR GO: GO:0007160; P: cell-matrix adhesion; TAS.
 DR GO: GO:0007517; P: muscle development; TAS.
 DR InterPro: IPR000413; Integrin_alpha.
 DR Pfam: PF01839; VWF_A.
 DR Pfam: PF00092; VWF_A.
 DR PRINTS: PRO1185; INTEGRIN_A.
 DR PRINTS: PRO0453; VWFADOMAIN.
 DR SMART: SM00191; Int_alpha; 5.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE: PS00234; VWA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 1189
 FT DOMAIN 23 1142
 FT TRANSMEM 1143 1165
 FT DOMAIN 1166 1189
 FT REPEAT 38 94
 FT REPEAT 102 163
 FT DOMAIN 187 345
 FT REPEAT 359 420
 FT REPEAT 422 475
 FT REPEAT 477 537
 FT REPEAT 539 598
 FT REPEAT 601 653
 FT DOMAIN 1154 1162
 FT DOMAIN 1174 1177
 FT CA_BIND 488 496
 FT CA_BIND 551 559
 FT CA_BIND 613 621
 FT DISULFID 76 83
 FT DISULFID 121 139
 FT DISULFID 139 139
 FT DISULFID 659 668
 FT DISULFID 674 729
 FT DISULFID 781 787
 FT DISULFID 881 893
 FT CARBOHYD 82 82
 FT CARBOHYD 95 95
 FT CARBOHYD 291 291
 FT CARBOHYD 331 331
 FT CARBOHYD 358 358
 FT CARBOHYD 449 449
 FT CARBOHYD 462 462
 FT CARBOHYD 528 528
 FT CARBOHYD 642 642
 FT CARBOHYD 694 694
 FT CARBOHYD 857 857
 FT CARBOHYD 894 894
 FT CARBOHYD 973 973
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1040 1040
 FT VARIANT 433 433
 FT VARIANT 524 524
 FT VARIANT 972 972
 FT VARIANT 1003 1003
 FT VARIANT 1030 1030
 FT VARIANT 1094 1094
 SQ SEQUENCE 1189 AA; 133609 MW; 6030309A4A4CD52 CRC64;

Query Match 18.6%; Score 1093.5; DB 1; Length 1189;
 Best local Similarity 28.3%; Pred. No. 1.1e-64;
 Matches 349; Conservative 214; Mismatches 502; Indels 167; Gaps 47;

QY 1 FNLDTENAMTFQENARG-FGQSVVQ--LQGR-VVVGAPQEI VAAANQSGSLYOCYSTGS 56
 Db 23 FNDMTAKRPVPGSRFAFTGTVVQHDHDSGNKVLVVGAPLETNGYKTDGVKCPVHGN 82
 QY 57 CEPRL-----QVPVEAVNMSLGLSAAATSPQLLACGPTVHQTCESENYKGLCLFLP 110
 Db 83 CTKLNLRVTLNVSERKONMRGLSLATNPKNKSNFLACSPFLMSHECGSSYITTCMSRV 142
 QY 111 GSNLRQOPKQFPAIRGCPQEDSDIAPLDGSGIIPHDERRMKEFVSTVMEQ--LKSK 168
 Db 143 NSNFRSKTVAP-ALQRC-QTYMDIVVLGDSNIYP--WVEVQHFLINILKFFYIPGQ 198
 QY 169 TLFSLAQYSEFRIHTFTEFQNNPNRSLVKPTQLLG-RTHPTATGVKRVIELNITN 227
 Db 199 IQGVGVQGEDVWHEPHLNDYRSVKDWEAAASHIEQGGTETRTAFGIEFARSEAFQ--K 256
 QY 228 GARKAKILIVITDGEKFGDPLGYEDVITPEADREGVIRYVIGV----GDAFSEKSRQ 282
 Db 257 GGRGAKKVMIVITDGHSDSP-DLEKVIQOQSRDNVIRYAVAVLGYNNRGINPFIPLN 315
 QY 283 ELNTIASKPPRDHVFQVNFPEALKTIONLREKIFAIETGTQSSSSPHEMSQEGPSAA 342
 Db 316 EIKYIASDPDDKHFFNVYDEAALKDIVDALGDRIFSLGNTK-NETSFGLMSQTFSSH 374
 QY 343 ITSNGPLISTVGSYDAGGVPLVTSKEK-----STPINMTRVDSMDNDAYLGYAAAILR 397
 Db 375 VVEDGVLLGAVGAYDMNGAVLKETSAKVIPLRESYLKPEELKXNHGAYLYTTSVVS 434
 QY 398 NRV-QSLVLGAPRYQHIHGLVAMP-RQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGS 455
 Db 435 SRQGVVAVAGAPRFNHTGKILFTMHNRSITIHQAMRGQIGSYFGSEITSVDIGDV 494
 QY 456 TDVLICAPHYEQTR-GGVSVCLPRQARWQCDVLYGEGQGWPRGALTVLGD 514
 Db 495 TDVLLGAPMYNEGRERGVVYVEL---RQNRVYVNGTLKDSHSYQNAQRFSSIASVRD 551
 QY 515 VNGDKLTDVAIGAPGEDNREGAVLPHGTSGSISPSHSORIASGLSPRLQYFGQSLSG 574
 Db 552 LQNSYNDVVVGAPLEDNHAQIYIFRGFRGS-ILKTPKQBITASELATGLQYFGCSIHG 610
 QY 575 QGDLTMDGLVDLTVGAQSHVLLASQVLRVKALMEFNPREVARNVP--EONDQVVGKGE 632
 Db 611 QLDLNEEDGLDLAVGALGNVILWSRPVQINASLHFEPSKI--NIFHRC-----KR 661
 QY 633 AGEVRVCL-----HVQKSTEDRLREQIQSVVTVYDLDLSDGRPHSAFVNET 679
 Db 662 SGRDATCLAAFLCTPTIFLAPHFQTTVG-----IRYNATMDERRYTPRAHLDEG 711
 QY 680 KNS-TRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLFVSGTPLSAFGLRPLVA 738
 Db 712 GDRFTNRAVLSSQGLCERINPHVLD-TADYVKPVTFSVEYSLEDP-----DHGPMLD 764
 QY 739 EDAQRLFTALPPEKNCNDNICODL-----SITPSF 771
 Db 765 DGWFTTLRVSVFPWNGCNEDESHCVFDLVDARSOLPTAMEYQORVLRKPAQDCSAYTLSP 824
 QY 772 MSLDCLVVGGRPREFNVTVVNDQEDSYRTQVTFPFDLDSYRKVSTLQNSORSWELA 831
 Db 825 DTYVFIESTFORVAVEATLENRGENAYSTVLNLSQSANLQP--ASLIQKEDSDGS--IE 880
 QY 832 CESASSTEVSGALSKTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENN 891
 Db 881 CVNEER-----RLQKQVCNVSYPFFRAKAKVAFELDFEP-SKSIPLHLELELAAGSDSN 934
 QY 892 MPRNTKTE--FQLELPVKYAVVTVVSHGVSTKY---LNFITAS--ENTSRVWQHQYQVSN 944
 Db 935 ERDSTKEDNVAFLPHLYKYEADVLFTRSSSLSHVEVKNLSLSEYVDGIGPPFCIFRQIN 994
 QY 945 LGQ---RSLPISLVPLVPLNQTIVWDRPQVTFSENLSSSTC----HTKELRPHSDFLA 997
 Db 995 LGLPPIHGMKTIPIATRSNRLKLRLDPLT-DEVANTSCNIWGNSTBYRPTPVE--E 1051
 QY 998 ELRKAPVNCISIAVCQRIQCDDIPFGIQEBEENATLKGNLSPDWY-----INTSHNLLIV 1052


```

Db 1052 DLRRAPQLNHSNDVVSINCINRLVP-NQEIFHLLGNL---WLSRLKALKYKSMKIWN 1107
Qy 1053 STAETLFDVDSVFTLLPGQAPVRSQTKVPPFVFN-----PLPLVGSVSGLLLLA 1106
Db 1108 AALGRQFH-SFF-----IFREDPSRQIVFISKQEDQVPIWIVGSTLGLLLA 1158
Qy 1107 LITLALYKLGPFK--RQYKMMSEGGPPGAP 1136
Db 1159 LVLALMKLGFGRARRR-----PGLDP 1183

```

RESULT 11

```

ITA2_BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
GN ITGA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94193647; PubMed=7511592;
RA Kamata T., Puzon W., Takada Y.;
RT "Identification of putative ligand binding sites within I domain of
RL integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
RL J. Biol. Chem. 269:9659-9663(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, INTEGRIN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC EMBL; L25886; AAB59255.1; -
CC PIR; I45914; I45914.
CC HSSP; P17301; LAOX.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC SMART; SM00191; Int_alpha; 5.
CC SMART; SM00327; VWA_1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER 1

```

```

FT SIGNAL 18
FT CHAIN 19
FT DOMAIN 19
FT TRANSMEM 1121
FT DOMAIN 1122
FT DOMAIN 1144
FT REPEAT 34
FT REPEAT 2
FT DOMAIN 177
FT REPEAT 2
FT REPEAT 423
FT REPEAT 477
FT REPEAT 540
FT REPEAT 604
FT CA_BIND 488
FT CA_BIND 552
FT CA_BIND 616
FT SITE 472
FT SITE 1146
FT DISULFID 72
FT DISULFID 669
FT DISULFID 778
FT DISULFID 854
FT DISULFID 1008
FT DISULFID 1044
FT CARBOHYD 94
FT CARBOHYD 101
FT CARBOHYD 332
FT CARBOHYD 421
FT CARBOHYD 449
FT CARBOHYD 464
FT CARBOHYD 688
FT CARBOHYD 748
FT CARBOHYD 945
FT CARBOHYD 1063
FT CARBOHYD 1070
FT VARIANT 580
FT VARIANT 588
FT VARIANT 725
SQ SEQUENCE 1170 AA; 128923 MW; ECEFC15F2448FB1 CRC64;

Query March 18.5%; Score 1084; DB 1; Length 1170;
Best local similarity 27.7%; Pred. No. 4.5e-64;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVQL---QGSRVVVGAPQEIIVAAVNRGSLYQC--DYST 54
Db 19 YNVGLPKAKTIPSGPSSEQFGVAVQOFINPKGNMLLVGSPWSPGPKNRMGDVYKCPDLST 78
Qy 55 GCEPIBLO-----VPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVVKGLC 107
Db 79 TTCCKLNLTSTMSNVTEMKTNMSLGLTLTRNVGTGFLTCGFLPNAQCCSQYTTGVC 138
Qy 108 FLPGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKPEFVSTVMEQLK-- 165
Db 139 SDVSPDP-QARTSPAPAVQTCF-SFIDVVVVCDSNSIYPWD--AVKNLEKFGVGLDIG 194
Qy 166 KSKTLFSLMYSSEFRHFTFKFQNNPNPSLVKPTQLL---GRHTATGVKRVIRE 221
Db 195 PTKTQMGLIQVANNPRVFNLTFRSKD---EMTKATSTQFYGGDLNTFKAIQADYT 251
Qy 222 LNIITNGARKNAFKLIVITDGEKFGDPLGYEDYVPEADREGVIRYIGV-----GDAFR 276
Db 252 AYSTAAGCPGATKVMVVVTVDGESH-DGSKLXAVIDCKNKLRFGLAVGLYLNRLD 310
Qy 277 SEKGRQELNTIASKPPRDHVFQVNRFEALKTIONLRKFIATICTQTGSSSSFEHNSQ 336
Db 311 TKNLIKETKAIASIPTRHFFNVSDADLEKAGTIGEQIFSIETVQG-GDNFQWMSQ 369
Qy 337 EGFSAAIT--SNGPLISTVGSVDNAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGVA 391
Db 370 VGFSAEYSPQNNILMILGAVGAYDSGTVQKTPHGLIFPSKQAFQIILQDRNMSYLGVS 429
Qy 392 AAIIILNRVQSLVLGAPRYQHIGLVAMPRQNTGMWESNANV-----KGTQIGVFGASL 445

```

```

Db 430 VASISTGNSVHFVACAPRANTYGCQVLYSVN-----ENGNTVTIQSQRGDQIGSYFGSVL 484
Qy 446 CSVDVDSMGSTDLVLIGAPHYEQTR--GGQSVSCPLPRGQARQACDAVLYGEGQCPWG 503
Db 485 CAVDVNKDITDVLVAGAPWYNDLKKEGRVYLFTIKG-ILNWH--QFLEGPNGLENA 541
Qy 504 REGAALTVLGVNGDKLTDVAICAPGEDNRCANVLFHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIALSDINMDGDFNDVIVGSPLENQNGAVYIYNGHEGM-IRLRYSKILGSDRAF 600
Qy 562 SPRLQVFGOSLGGQDLTWGDLVDTVGAQGHVLLRSQPVLRKVAIMEFNPREVARNVF 621
Db 601 SHELQYFESLGYDGLNGDSITDVSVCAGQVQVQLMSQSIADVSVDASFTPKXI--TLL 658
Qy 622 ECDQVVGKGEAGEVRVCLHVQKSTRDLRREGIQSVVTVYDLAD-----SORPHSAVFVN 677
Db 659 NKNABSI-----KLKLCF---SAKPRFTQNNQVAIVYITIDQEDFSRVSIRGLEK 707
Qy 678 EFGNSTRTQVGLGTQCE--TLKQLPNCIEDPVSIVLRLEFSL--VGTPLSAFGLN 733
Db 708 ENNERCLKTHIVSAQRCSEVYIIHQEPS-----DIISPLNLCMISLENPGT----- 756
Qy 734 RPLVAEDAQRLTALFPPEKNGCNDNI CDDLSITF-----SPMSLDCLVWGGPREFNTV 789
Db 757 NPALRAYSETVAFSIPFKHKGDDGVCISDLVLNVQQLPATQQQPFIVSNQNKELTFSV 816
Qy 790 FVRNDGEDSYRTQVTFEPFLDLSYRKVSTLQNRQSRWRACESASST-EVSGALKSTS 848
Db 817 QLNKKEGAYNTEIVVDSENLFF-----ASKMPVDGTEVTCQIASQKSVT 864
Qy 849 CSINHPITPENSEVFTNTPDVSXASLGNKLLKANVTSENMPRTNKTFFQLPVPKY 908
Db 865 CNVGYPALKSQKVTFITNFDNLQ-NLQNASISFRALSSEQENMADNSVNLKLSILY 923
Qy 903 AYVMVTSHGVTKYLNTASNTSRVQHQYQVSNLQCR-----SLPISLFLV 958
Db 924 DAEIHIT-RSTNINPYEVLSDGNVSSV-HSPE--DIGKPFISIKVTGVSVPVSM- 976
Qy 959 PVRLNQTVWRPQVTFSEN---LSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVLIHPQYTKDKNPLWLTGVHTDQAGDISCEABINPLKIGQTSVSVFSKEN 1030
Qy 999 LRKAPVNCISIAVCORIQQDIPFGIQIEFNATLKNLSFDWYIKTSNHLIIVSTAEI- 1057
Db 1031 FFIKELNCRATSCSNIMCWLESDLVQKGEYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNDVSFTL-----LPGQAFVRQETETKVEPFE-VENPLPLIVGSSVGGILLIALITA 1110
Db 1091 TYNPQIYVIEENTVIP-----LTINKPHEKVEVPTGVIVGVIAGILLALLVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFRRKRYERM 1155

RESULT 12
ITR2_MOUSE
ID ITR2_MOUSE STANDARD; PRT: 1178 AA.
AC Q62459; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CSA; TISSUE=Lung;

```

```

RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Danjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RL not virus binding.";
RN Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.B., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RN Dev. Dyn. 199:292-314(1994).
RC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z29987; CAAB2877.1; -.
DR EMBL; X75427; CAAS3178.1; -.
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
RW Platelet; Signal; Repeat; Calcium; Magnesium.
RW SIGNAL 1 26
RW CHAIN 27 1178
RW DOMAIN 27 1129
RW TRANSMEM 1130 1151
RW DOMAIN 1152 1178
RW REPEAT 42 100
RW REPEAT ? ?
RW DOMAIN 185 375
RW REPEAT ? ?
RW REPEAT 431 483
RW REPEAT 485 546
RW REPEAT 548 607
RW REPEAT 612 664
RW REPEAT 496 504
RW CA_BIND 560 568
RW CA_BIND 624 632
RW SITE 480 482
RW SITE 1154 1158
RW DISULFID 80 89
RW BY SIMILARITY.
RW CELL ATTACHMENT SITE (POTENTIAL).
RW GFGR MOTIF.
RW BY SIMILARITY.

```

FT DISULFID 677 734 BY SIMILARITY.
 FT DISULFID 786 792 BY SIMILARITY.
 FT DISULFID 862 873 BY SIMILARITY.
 FT DISULFID 1016 1047 BY SIMILARITY.
 FT DISULFID 1052 1057 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 18.2%; Score 1072; DB 1; Length 1178;
 Best Local Similarity 28.1%; Pred. No. 2, 9e-63;
 Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

QY 1 FNLDTNAMTFO--ENARGGQSVQI---CGSRVVVGAFOEIVAAQNRGLQC--DYST 54
 DB 27 YNVLFGAKIFGSPSEQEGYSVQVLTNPQGNWLLVGPSPGFPEMRMGDVYKCPVDLPT 86
 QY 55 GSCERLQO-----VPVEAVMSLGLSLAATSPQLLAGPTVHQTCSENTYVKGLC 107
 DB 87 ATCEKANLQNSASISNTEIKTMSLGLTLTENPGTGGFLTQPLWHQCGNQYATGIC 146
 QY 108 FLFGSLRQOPQ---XFPALRGCPQSDPIALIDGSGIIPHDPFRMKFVSTVMEQL 164
 DB 147 ---SDVSPDFQPLTSFSAVQACPSL--VDVVVVCDESNIYP--WEAVKNFLVKPVTGL 199
 QY 165 K--KSTLSLMOYSEPHIHTFKFQONNPRSLVKPITQLLG--RTHATGVRKVIRE 221
 DB 200 DTGPKKTQVALIQYANEPHIIFNLDFTFKQMVQATSETROGGDLTNTFRAIFARDY 259
 QY 222 LLNITNGAKNAPKILIVITDGEKFGDPLGYEDVIPEADREGVIRYVIGV-----GDAFR 276
 DB 260 AYSQTSGGPRGATKVMVVYTGESH--DGSKLTVIQCCNDDEILRPGIAGVLGNALD 318
 QY 277 SEKSRELNTIASKPRDRHVQNNFEALKTQNRKIPAEIGTQTCSSSPFEHXSQ 336
 DB 319 TKNLKEIKAIATSTPTERYFFNVADBAALLEKAGTLGQIFSGIEGTVOG--GDNFQKEXAQ 377
 QY 337 EGPSA--AITSNGPLASTVGSYDWAGGVFLYTSKEXSTFINMT--RVDSDMN--DAYLGYA 391
 DB 378 VGFSAADYARQNDILMLGAVGAPDWSGTIVQETSHKXVIIPKQAFDQVLDRNHSFLGYS 437
 QY 392 AAILIRNRVQSLVGLAPRYOHIGLVAMFRQNTGWSNANV-----XGTOIGAYFGASLCS 447
 DB 438 VAAISTEDGVHPVAGAPRANYTCQIVLYSVNK---QGNVTVIQSHREGDQIGSYFGSVLCS 494
 QY 448 VDVSNGSTDLVLIGAPHYEQTR--GGQVSVCPLRGQARQCDAVLVGRQGPWGRF 505
 DB 495 VDVKDQITIDVLLVGAPTWNLDKBEGRVYLFITTKGILNHQ----FLEGPEGTGNARF 551
 QY 506 GAALTVLGDVNGDKLTDVAIGPEEDNRGAVVLPHTGTSGETSPSHSQRIAGSKLSPR- 564
 DB 552 GSAIAALSDINMDGENDVIUGSVENENSGAVVYNGHQT--IRTKYSQKILGNSGAPR 610
 QY 565 -LQYFGQSLSCGQDLTDGLVDITVGAQGVLLRSQPLVRKAIEMFNPREVARNVFEC 623
 DB 611 HLQFFGRSLGVDLMDGSDITDYSIGALGQVQLWSQSIADVAIEALFTP----- 660
 QY 624 NDQWVGKAGEVRVCLHVQKSRDLRREGIQSVVTVYDALD----SCRPHSRVAFNET 679
 DB 661 -DKITLLNKDAKITLKCFAEFAPAGQNNQV--AILFNMTLDADGSSSRVTSRGVREN 717
 QY 680 KNSTRQTVGLTQTCET--LKLQLPNCIEDVPSIVIRLNPSPVGLTSLFAGNLRPVL 737
 DB 718 SERFLOKVMYVNEVQKSEHHISIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
 QY 738 AEDAQRLLFTALFFPEKXGNDNQCQDLSI-----TFSPMSLDCVLVGVGPREFNTV 788

DB 769 EAYSETVKVFGSIPFYKCGSDGICISDILDVQQLPAIQTSF-----IVSNQNRKLRFTS 823
 QY 789 VTVNDGDSVRTQVTPFPPLDLGYRKVYSTLQNRQSQRSLACESASST--EVSGALKST 847
 DB 824 VILKNRGESANTVTVLAFFSENLF-----ASPSMPVDGTEVTCVGSQSKSV 871
 QY 848 SCSINHPIPPENSEVETNITFDVDSKASLGNKLLKANVTSENMPRTNKE--PQLBPL 905
 DB 872 TCDVGYPAKSEQQVTFITNFDFNLQ--NLQQAALNFQAFSESQ--ETNKADNSVSLTIP 928
 QY 906 VKYAVVMVVTSHGVSTKVLNFTASENTSRVMQHVQVSNLQOR-----SLPISLV 955
 DB 929 LQYDAELALT--RSTNINFEIISSENAPSVIK--SVEDIGPKFIPSLKVTAGSAPVMA 984
 QY 956 FLV-----PVRNLQTVIWDPRQVTP--SEMLS 980
 DB 985 LVTHIHIPQYTKENPELLYLTGTDQAGDISCTAEINPLKLPHTA----PSVSFKENFR 1040
 QY 981 STCHKEELPSHSDPLAELRKAPVVCNSIAVCORIQCDIPFGIGQEBFNATLKGNSLFDW 1040
 DB 1041 ---HTKE-----IDCRTTSCSNITCWKLDLHMKAEYFINVTTVVNR 1080
 QY 1041 YIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTTETKVEPEVFNPLPLIVGSSVG 1100
 DB 1081 FAASTFQVQVLTAAAEIDTHNPQLFVIBENAVTIPLMTMKTEKAEVPT--GVIIGSIIA 1138
 QY 1101 GILLALITALYKLGFPKROYKDM 1125
 DB 1139 GILLALLAMTAGLWKLGLFFKROYKDM 1163

RESULT 13
 ITA2 HUMAN
 ID ITA2_HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN ITGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
 RC TISSUE-Endothelial cells;
 RX MEDLINE=89308879; PubMed=2545729;
 RA Takada Y., Hemler M.E.;
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit (platelet GP1a): homology to other integrins and the presence of a possible collagen-binding domain.";
 RL J. Cell Biol. 109:397-407 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
 RX MEDLINE=98019223; PubMed=9353312;
 RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;
 RT "Crystal structure of the I domain from integrin alpha2beta1.";
 RL J. Biol. Chem. 272:28512-28517 (1997).
 RN [4]
 RP VARIANT HPA-5 (BR).
 RX MEDLINE=94043762; PubMed=7901236;
 RA Santoso S., Kaib R., Walka M., Kiefel V., Mueller-Eckhardt C.,
 RA Newman P.J.;
 RT "The human platelet alloantigens Br(a) and Brb are associated with a single amino acid polymorphism on glycoprotein Ia (integrin subunit

FT	REPEAT	45	103	FG-GAP 1.
FT	REPEAT	188	378	FG-GAP 2.
FT	DOMAIN	188	378	WVFA.
FT	REPEAT	433	433	FG-GAP 3.
FT	REPEAT	434	486	FG-GAP 4.
FT	REPEAT	488	549	FG-GAP 5.
FT	REPEAT	551	610	FG-GAP 6.
FT	REPEAT	615	667	FG-GAP 7.
FT	CA_BIND	499	507	POTENTIAL.
FT	CA_BIND	563	571	POTENTIAL.
FT	CA_BIND	627	635	POTENTIAL.
FT	SITE	1157	1161	GFPR MOTIF.
FT	DISULFID	83	92	BY SIMILARITY.
FT	DISULFID	680	737	BY SIMILARITY.
FT	DISULFID	789	795	BY SIMILARITY.
FT	DISULFID	865	876	BY SIMILARITY.
FT	DISULFID	1019	1050	BY SIMILARITY.
FT	DISULFID	1055	1060	BY SIMILARITY.
FT	CARBOHYD	105	105	N-LINKED (GLCNAC. .)
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. .)
FT	CARBOHYD	343	343	N-LINKED (GLCNAC. .)
FT	CARBOHYD	432	432	N-LINKED (GLCNAC. .)
FT	CARBOHYD	460	460	N-LINKED (GLCNAC. .)
FT	CARBOHYD	475	475	N-LINKED (GLCNAC. .)
FT	CARBOHYD	699	699	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1057	1057	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1074	1074	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1081	1081	N-LINKED (GLCNAC. .)
FT	VARIANT	534	534	X -> E (IN ALLOANTIGEN HPA-5B; dbSNP:1801106) /FTID=VAR_003977.
FT	TURN	170	171	
FT	STRAND	173	180	
FT	TURN	183	184	
FT	HELIX	188	199	
FT	TURN	200	201	
FT	STRAND	204	204	
FT	TURN	206	207	
FT	STRAND	209	216	
FT	STRAND	220	224	
FT	TURN	226	228	
FT	HELIX	232	240	
FT	TURN	241	241	
FT	HELIX	252	262	
FT	HELIX	263	264	
FT	HELIX	266	268	
FT	TURN	269	269	
FT	STRAND	275	282	
FT	HELIX	289	291	
FT	HELIX	292	301	
FT	TURN	302	303	
FT	STRAND	304	311	
FT	HELIX	313	317	
FT	TURN	318	319	
FT	HELIX	323	330	
FT	TURN	331	332	
FT	HELIX	337	340	
FT	STRAND	341	344	
FT	HELIX	347	353	
FT	HELIX	354	362	
FT	TURN	363	363	
SO	SEQUENCE	1181 AA;	129295 MW;	7E1B7ED968A94070 CRC64;

Query Match 18.2%; Score 1069; DB 1; Length 1181;
Best Local Similarity 27.0%; Pred. No. 4.6e-63;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43

Qy	1	FNLTENAMTFQ-ZNARGFGOSVVLQ---	QGSRVVWGAPOEIVAANQSGSLYOC--DYST 54
Db	30	YNYGLPEAKIFSGPSSRFQGVAVQCFINPKENMLLVGSPWSGFFPENRMGDVYKCPVDLST 89	
Qy	55	GSCEPISLQ-----VPVEAVNMSLGLSLAATTSPPLLACGPTVHOTCSENTYVKGLC 107	

Db 90 ATCEKLNQTSSTSPNVTETKMSLGLILTRNMGTCGLTCPLMAOCCNQYITTVGC 149
Qy 108 FLPGNLRQOKPPEALRGCPQSDSDIAFLIDSGSHIIPHDPRMKEPFTVWQELK-- 165
Db 150 SDISPDF-QLSASPSAPQPCPSL-IDVVVVCDSNSIYPWD--AVKNLEKFKVQGLDIG 205
Qy 166 KSKTLFSLMOYSBBERIHFTEFKBQNNPNRSLVKPTLQLLG-RTHTATGVKRVIRELLN 224
Db 206 PTKQVGLIQAANPRVFNLTNTYKTEEMIVASOTSOYGGDLTNFTGAIQYARKAYS 265
Qy 225 ITGARKNAFKLIVITDEKFGDPLGYEDVIPADREGVIRYIVGV-----GDAFRSEK 279
Db 266 AASGRRRATKVMVVTVDGSH-DGSMLEKAVITDQCNHNLIRFGIYVLYNLRNALTQKN 324
Qy 280 SRQELNTASPPRDHVFQVNFPAKTIQOLAREKIPAEISTGTGSSSSPEHMSQEGF 339
Db 325 LIKEIKATASIPTEFYFNVSDEALKEAGTLGEGHFIESTVQG-CDNFQMSQVGP 383
Qy 340 SAAITSNGP--LLSTVGSYDMAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYIGYAAAI 394
Db 384 SADYSSQNDILMLGAVGAGFGSGTIVOKTSHGLIFPKQAFQDQILODRNHSYLGYSVAA 443
Qy 395 ILNRNVQSLVJGAPRYQHILGVAMFRONTGWNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGSETHFVAGAPRANTYQGVLYSVN-----ENGNITVTOAHRGQDQIGSYFGSLCSV 498
Qy 449 DVDNSGTDVLIGAPHYVEQTR--GGQVSVCPLEPRGQARWQCDVAVLYGEGQFQWGRFG 506
Db 499 DVDKDTITDVLVAGPMYSDLKKEEGRVYLTIKKGILGQH-----FLEGPEGIENTRFG 555
Qy 507 AALTVDGVNGDKLTDAIGAPGEDNRGANVLFEGTSGSGISPSHSORIASG--KLSPR 564
Db 556 SAIAALSINMDGDFNDVIGSPLENQNSGAVIYNGHQT-IRTKYSQKILGSDGAPRSH 614
Qy 565 LOYFGQSLGGDLTMDGLVDLTVCAQGHVLLLRQPVLRKVAIMEFNPREVARNVFCN 624
Db 615 LYIFGRSLDGYDLNGSDSTVDSIGAFQGVQLMSQSIADVAIEASFPEKI--TLVNKN 672
Qy 625 DQVWKGAGEVRVCLHVOKSTRDLREGQISQVTVYDLAD-----SGRPSRAVFNETH 680
Db 673 AQII-----LKLCP-----SAKFRPTKQNNQVAIVNITLDADGFSRVSRTSGFLKKN 721
Qy 681 NSTRACTQVLGLTQTC--ETLKLQPLNCIEDPVSPIVLRNLSLVTPLSAPGNLRPVLA 738
Db 722 ERLQKQVMVNAQSCPEHIIYQEPS-----DVNSLDLRVDSLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFFPKNCNDNICODLSITP-----SPMSLCLVGVGPRFNVTVVRND 794
Db 773 AYSETAKVSIIPFKDCGDLGCLSDLVLDVRAQIPAAQOPRIVSNQNKRLTFSVTLNKK 832
Qy 795 GEDSVRTQVTFPPLDLSTRKYSTLQNSQSRWLACESASST-EVSCALKSTSCSINH 853
Db 833 RESAYNTGIVDFSENLPF-----ASFSLPVDGTEVTCQVAASQKSVACDVGT 880
Qy 854 PIPPENSEVTNITFDVDSKASIGENKLLKANVTSENMPRTNKTETFOLELPVKYAVVMV 913
Db 881 PALKREQVTFINFDNLQ-NLQNCASLSFQALSESQENKADNLVNLKILLYDAEI- 938
Qy 914 VTSHGVTYKLYNAPTASENTSRVMOHYQVSNLQOR-----SLPSLVELV----- 958
Db 939 ---HLTRSTNINPFISSDGNVSIHVHSEFVGEKFIPLSKVTTGSPVPSMATVVIHIPQ 995
Qy 959 -----PVLNQVTVIWRDPOVTP-SENLSSTCHTKER 988
Db 996 YTKERPLMYLTVQTDKAGDISCNADINPLKIGQT-----SSSVSPKSENER---HTKE- 1047
Qy 989 LPSSHDFLAELRKAPVNVNCSIAVCORIQCDIPFGIOEFNATLKNLSDFWYVITSNH 1048
Db 1048 -----LNCKRTASCSNVTCWLDKDVHMKGEYFVNVTVIIRIWNGTFFASSTFQT 1091
Qy 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRSQTEKTVBEFVENP-----LP--LIVGSSVG 1100
Db 1092 VQLTAAAEINTNPFIYVI-----EDNTVTIPLIMKDEKAEVETGVIGSIITA 1141

Qy 1101 GILLALITAAALYKLGFFKROYKDM 1125
Db 1142 GILLALVAILWKLGGFFKKYKEM 1166
RESULT 14
ITAG HUMAN
ID ITAG HUMAN STANDARD; PRT; 1167 AA.
AC 075578; Q9UH28;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Akerlund E.,
RT "Isolation, cloning, and sequence analysis of the integrin subunit
RT alpha10, a beta1-associated collagen binding integrin expressed on
RT chondrocytes.";
RL J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
RA Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
RT structure, and chromosomal localization.";
RL Cytogenet. Cell Genet. 87:238-244(1999).
CC 1- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC 1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC ASSOCIATES WITH BETA-1.
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- TISSUE SPECIFICITY: Widely expressed with highest expression in
CC muscle and heart. Found in articular cartilage.
CC 1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC 1- SIMILARITY: Belongs to the integrin alpha chain family.
CC 1- SIMILARITY: Contains 1 VWFA domain.
CC 1- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF074015; AAC31952.1; -;
DR EMBL; AF112345; AAF21944.1; -;
DR EMBL; AF172723; AAF61638.1; -;
DR HSSP; P17301; IAOX.
DR Genew; HGNC:6135; ITGA10.
DR MIM; 604042; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1123 1145 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1146 1167 POTENTIAL.
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT 38 97 FG-GAP 2.
 FT REPEAT 38 97 VWA.
 FT DOMAIN 167 350 FG-GAP 3.
 FT REPEAT 355 427 FG-GAP 4.
 FT REPEAT 428 482 FG-GAP 5.
 FT REPEAT 483 545 FG-GAP 6.
 FT REPEAT 546 605 FG-GAP 7.
 FT REPEAT 608 660 POLY-LEU.
 FT DOMAIN 1134 1140 POTENTIAL.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 BY SIMILARITY.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 799 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 I -> L (IN REF. 2).
 FT CONFLICT 844 844 G -> V (IN REF. 2).
 FT CONFLICT 909 909 E -> D (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; A67D3A1C25C1ABAO CRC64;

Query Match
 Best Local Similarity 18.1%; Score 1065.5; DB 1; Length 1167;
 Matches 358; Conservative 196; Mismatches 488; Indels 181; Gaps 46;
 1 FNLDENAMTFQENARG-FQGVVQLQGRS---VVVGAPOEIVAAQNGSLYQC----- 50
 23 FNLDENHPRLPFGPPAEFGYSLQHVGGQRMVAVGAPWDGPGSDRGDVRCPVGGAH 82
 51 -----DYSTG-SCEPIRLQVPEAVNMSLGLSLAATSPPOLLAGGTVHTQCS 99
 83 NAPCAKGLRGVQLGNSSHP-----AVNHLGMSLLETGDGGFMACAPLWSRAGS 134
 100 NTYVKGGLCFPGSNLRQQPKPEALRGCPQEDSDIAPLDGSGSIIPHD-----PRRM- 153
 135 SVFSSGICARVDASFPQGSLATQOR-CETY-MDVVIVLDGNSIYPMSEVQTFLLRLV 192
 154 -KEFVSTVMEQLKKSKTLFSLMYSREFFRIHFTFKFQNNPNRSLVKPITOLLGR-THT 211
 193 GKLFDIP--EQIQ-----VGLVQYGPSPVHVESLGDFTKKEVRAAKNLSRREGRETKT 245
 212 ATGVKRVIRELLNITGARKNAFKILIVITDGKFF-GDPLGYEDVPEADREGVIRYVIG 270
 246 AQAIMVACTEGFSQSHGGRPEARLLVVDVGDHGEELPAALACEAGR--VTRYGIA 303
 271 V-GDAFSEKSS---RQELNTIASKPPDRHVFQVNNFEALKTIQNLQREKIFAIBGTQFG 325
 304 VLGHYLRQRQDPSPFLREIRTIASDPDRFPFVNTDEAALTDIVDALGDRIFGLESHAE 363
 326 SSSSFEHMSQEFSAITSNGLLSTVGSYDHAGGVFLYTSKEKSTFFINWTVDS----- 381

364 NESSFGLMSQIGFSTHRLKDGLFGVAVGAYDWGSGVLMLEGGHRLFPFPRMALEDFPPA 423
 382 -DNNDAYLGYA-AAIILNENRQSLVLGAPRYQHIGLVAMFR-ONTGWNESNANVKGQIG 438
 424 LQNHAYLYGYSVSSMLRGGRRFLFSGAPFRGKVIAPOLKDGAVRVAQSLQGGQIG 483
 439 AYFGASLCSDVDNSGSDTLVLGAPHY--EOTRGQVSVCPLEPRQORARWQCDVLYG 496
 484 SYFGSELPLDTRDGTDTLLVAAPWFLGPONKGTCTRVVYLV--GQSSLLTLQGTLP 541
 497 EOGQWGRFGAALTIVLDVNGDKLTVAIGAPGEENRGAVLFGHSGSGISPSHSORI 556
 542 EPPQD-ARFGFAMGALPDNLQDGFADVAVGAPLEDGQGALYLYHGTO-SGVAPHAPRI 599
 557 AGSKLSPRLQYFQGSLSGGQDLTMDGLDVLTVGAQGEVLLLRSPQLRVKAIMFNPREV 616
 600 AASMPHALSYFGRSVDRLDGLDDLDVAVGAQGAAILLSRRIVHLTPSLEVTPOAI 659
 617 ARNVFENDQVVKKEAG--EYVAVCLHVQSTDRLRREGQIOSWVYDIALDSCGRHSA 674
 660 SVVQRDCRR--RQGEAVCLTAALCFQVTSRTGMDH---QFYMRFTASDEWTAGARA 713
 675 VFNET--KNSTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNLPFLVGTPLSPGN 732
 714 AFDGSGQRLSPRLRLSVG-NVICEQLHFHVL-D-TSDYLRPVALTVTFALDNTTKPG-- 768
 733 LRPVLAEDAQRLFTALPPPKNGNDNMCODLSITFSMSLDC-----LVVGGPRE 784
 769 --PVLNBSPTSIOKLVPFSKDCGPDNECVTLVQ--VNMDIRGSRKAPFVVRGGRKK 823
 785 FNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWFLACESASSTEVSCAL 844
 824 VLVSTTLNKENAYNTLSLISRNL--HLASLTQR-ESPIKVECAAPSA----- 872
 845 KSTSCSINHPIPPENSEVTNITPDVDSKASLG---NKLL-----LKANVTSENMPRTN 896
 873 HARLCSVGHVPVQTAQVTLLEPEFSCSLLSQVPGKLTASSDSLRNGTLQNTAQT- 931
 897 KTEQLZLPVYAVYVMTSHGVSTKYNLFTASNTSRVMOHY-----Q 941
 932 -----SAYIQVEPH-----LFSSESTLHYEVHPVGTLPVGPPEKTLR 973
 942 VSNLG---QRSPLISLVLPV-----VRLNQTVMRDPQVTFSENLSTCHTKERL 989
 974 VQNLGCVVSGLIISA--LLPAVAHGNYFLLSQVI-----TNASCIVQNLTEP 1022
 990 PSHGDFLAELRKAPVNVCSIAVCQICDIPFGIQEFNATLKNLSFDWVYKTSNHL 1049
 1023 PGPPVHPEELQHTNRLNGSNTQCVVRCHLQKLGTEVSGLLRLVINEFFRAKPKSL 1082
 1050 LIVSTAEILFNDVFTLLPGQAFVRSQTEKTVKFFVEPNELPLIVGSSVGLLILALIT 1109
 1083 TVSTFELGTEGSLVQLTEASRWSLLEV-VQTRPILISLWILIGSLVGLLILALLV 1141
 1110 AALYKLGFF-----KRYK 1123
 1142 FCLWKLGFPAHKKIPEEKREEK 1164
 RESULT 15
 ID ITAL RAT STANDARD; PRT; 1180 AA.
 AC P18614;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1) (CD49a).
 GN ITGAI.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

Db 621 PSGBGKTLKFRGQSIHGMXNDGLTDTVTIGGLGGAALFWARDVAVVVKVTMNFENKV 680

QY 617 ARNVFECNDQVVKGEAG--EVRVCLHVO-KSTRDLREGQIQSVVTVYDLALDSGRPHSR 673

Db 681 NIQKNCR---VEGKETVCINATMCFHVXKSKEDSIYEADLQ-----YRVTLDSLROISR 733

QY 674 AVFNER-----KSTRRTQTVGLTQTCETLKLQLPNCI-----EDPVSPIVLR 718

Db 734 SPFSGQERKIORNITVRESE-----CIRHSFYNLKDHDFQDSVRVTL 776

QY 719 NFSLVGTPLSAFGNLPVLAEDAQRFLTALPFPEKNCNDNICODDLSITFSFNSLDCIV 778

Db 777 DENLT-DPENG-----PVLDDALPNSVHEHIFPAKDCGNKERCISDLTLNVSTTEKSILI 830

QY 779 VGGPRE-FNVTVVRNDGSDSYRTQVTFPEPDLDSYRKVSTLQONQSSQBSWELACESASS 837

Db 831 VRSQHDKFVSLTVKXKGSAYNTRTVQVQSPNLIFSGIEEIQD-----SCEN-- 880

QY 838 TEVSGALKSTCSINHPIFPENSEVFNITFDVDSKASLGNKLL-LKANVTSENNMPTN 896

Db 881 -----QNTCVGVYFFRAGETVTFKILFQFNTSHLSENALIHLSATSDSEEPLESIN 933

QY 897 KTEFQLELFPVKXAV-----XNVVTSBGVST-----KXNFTASNTSRVMOHQYQVSNL 945

Db 934 DNEVNIPIPVKYEVLQFYSSASEHHISVAANETIPEFINST--EDIGNEINNVFTIRK 991

QY 946 QOESLP--ISLVP-----LVPVRLNQTIVD-----RP-----Q 972

Db 992 GHFPMPELQISIFPNLTADGYPVLPIG-----WSSDNVNCPRSELEDPFGINSKX 1045

QY 973 VTFPS-----ENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQIQCDI--- 1019

Db 1046 MUISKEVLKRGTIQDCSTC-----GVATITCSLLPSDLSQVNVSL 1088

QY 1020 ---PFTGIQEBP---NATLXGNLSPDWYIKTSHNHLLIYSTAEILLFNDVSFTLLPCQAF 1073

Db 1089 LKKPTF-IRAHFSSLNLTARGELK-----SENSLTJSSN----- 1123

QY 1074 VRSOTETKVEPFEVNPPLPL--IVGSSVGLLILALITAALYKLGFPFKQYKMMSE 1128

Db 1124 RXRELAIQISKDGLPGRVPLWILLSAPAGLELLMLLILALWKIGFPFKPLKKOMEK 1180

Search completed: June 7, 2004, 17:13:03
Job time : 13.9719 secs

QY 1 FNLDTENAMTFQENARFGQSVVQLQSGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQSGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLGSLNRQPOK 120
DB 77 PLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLGSLNRQPOK 136
QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTTFEQQNPFRSIAKPEITOLLGRTHATGVRKVIKRELMITNGARKNAFKILIVI 240
DB 197 RIHFTTFEQQNPFRSIAKPEITOLLGRTHATGVRKVIKRELMITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKEPRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKEPRDHVFOVN 316
QY 301 NFEALKTIONLRKPIFAIGTGTGSSSSSEHMSQGFSAITNSGFLSTGVSQDWAG 360
DB 317 NFEALKTIONLRKPIFAIGTGTGSSSSSEHMSQGFSAITNSGFLSTGVSQDWAG 376
QY 361 GVPLVTSKESPFINMTVDSDMDAYLGXAAIILNRVQSLVLAGPRYOHIGLVAWFR 420
DB 377 GAFLYPSKDKKAFINTTRIDSDMDAYLGXAAIILNRVQSLVLAGPRYOHIGLVAWFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDMSNGSTDVLVLGAPHYBQTRGGQSVVCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDMSNGSTDVLVLGAPHYBQTRGGQSVVCP 496
QY 481 PRGQBARQCDVLYGEQOPWGRFGAALTVDVNGDKLTDVAIGARBEEDNCGAVLF 540
DB 497 PRGQBARQCDVLYGEQOPWGRFGAALTVDVNGDKLTDVAIGARBEEDNCGAVLF 555
QY 541 HGTSGSGISPSHRSORIASGKSPRIQYFGQSLGSGQDITMDGLVLTGACQHVLLLSAQ 600
DB 556 HGTSGSGISPSHRSORIASGKSPRIQYFGQSLGSGQDITMDGLVLTGACQHVLLLSAQ 615
QY 601 PVLRVKALMEFNPRVARNVPCNDQVVKGEKAGEVRVCLLVQKSTRRLBEGQIQSVTT 660
DB 616 PVLRVKALMEFNPRVARNVPCNDQVVKGEKAGEVRVCLLVQKSTRRLBEGQIQSVTT 675
QY 661 YDLALDSGRHRAVNETNSTRTQTOVLGTCTKQLPNCIEDPVSPIVLRNF 720
DB 676 YDLALDSGRHRAVNETNSTRTQTOVLGTCTKQLPNCIEDPVSPIVLRNF 735
QY 721 SLVGTPLSAFGLNLRVLAEDAQRLFTALPPPEKNCNDNIODDLSITFSWMSDCLVVG 780
DB 736 SLVGTPLSAFGLNLRVLAEDAQRLFTALPPPEKNCNDNIODDLSITFSWMSDCLVVG 795
QY 781 GPRENVTVTVNDGEDSVRTOVTFPFLDLISYKVTILQNSORSNRLACESASSTEV 840
DB 796 GPRENVTVTVNDGEDSVRTOVTFPFLDLISYKVTILQNSORSNRLACESASSTEV 854
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRNTKTEF 900
DB 855 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRNTKTEF 914
QY 901 QLELPKVIYVWTVSHGVTKLNFASNTSRVMOHQYQVNSLQBSLPLSLVFLVPV 960
DB 915 QLELPKVIYVWTVSHGVTKLNFASNTSRVMOHQYQVNSLQBSLPLSLVFLVPV 974
QY 961 RLKQTVIWRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVVNCISVACQRIQCDIP 1020
DB 975 RLKQTVIWRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVVNCISVACQRIQCDIP 1034
QY 1021 PFGIQEENATUKGNLSPWYIKTSHNHLIIVSTABILPNDVFTLLPQGAIVFESQET 1080
DB 1035 PFGIQEENATUKGNLSPWYIKTSHNHLIIVSTABILPNDVFTLLPQGAIVFESQET 1094
QY 1081 KVEPPEVNPPLIVGSSVGLLILLALITAAALKYKLGFFKQYKQKQWMSGEGPGEAPQ 1137

DB 1095 KVEPPEVNPPLIVGSSVGLLILLALITAAALKYKLGFFKQYKQKQWMSGEGPGEAPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
AC Q8CA73
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR F730045J24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=2234683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK039444; BAC30350.1; -
DR PIR; PT0572;
DR PIR; PT0633; PT0697;
DR PIR; PT0697; PT0697;
DR MED; MGI:96607; Itgam.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWPA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;
Query Match 66.5%; Score 3907.5; DB 11; Length 1036;
Best Local Similarity 66.4%; Pred. No. 3.2e-276;
Matches 756; Conservative 121; Mismatches 142; Indels 119; Gaps 2;
QY 1 FNLDTENAMTFQENARFGQSVVQLQSGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQSGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLGSLNRQPOK 120
DB 77 PLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLGSLNRQPOK 136
QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFFVSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTTFEQQNPFRSIAKPEITOLLGRTHATGVRKVIKRELMITNGARKNAFKILIVI 240
DB 197 RIHFTTFEQQNPFRSIAKPEITOLLGRTHATGVRKVIKRELMITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKEPRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKEPRDHVFOVN 316
QY 301 NFEALKTIONLRKPIFAIGTGTGSSSSSEHMSQGFSAITNSGFLSTGVSQDWAG 360
DB 317 NFEALKTIONLRKPIFAIGTGTGSSSSSEHMSQGFSAITNSGFLSTGVSQDWAG 376

```
Qy 361 GVFLYTSKEKSTFTINMTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVPL 480
Db 437 ENPTWEPHTSING----- 450
Qy 481 PRGORARWQCDVAVLYGQGPWGRFGAALTVLGVDVNGKLTVDVAIGAPGBEDNRGAVILF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHSORLAKSLSPRLQYFGOSLGGQDITMDGLVDLTVGAQGHVLLARSQ 600
Db 451 -----SORIIGAHPSPQLQYFGOSLGGQDITMDGLVDLTVGAQGHVLLARSQ 498
Qy 601 PVLKVKAIMENPREVARNPECDQVVKKGAGEVRVCLVQKSTRDRLAEGIOQSVVT 660
Db 499 PVLKLEATMEFSPKKVARSVFACQVQVKNKADAGEVRVCLVQKSTRDRLAEGIOQSVVT 558
Qy 661 YDLALDGRPHSRVAFNETKSTRTOVGLTQTCETLKQLPNCIEDPVSPIVLBNF 720
Db 559 YDLALDGRPHSRVAFNETKSTRTOVGLTQTCETLKQLPNCIEDPVSPIVLBNF 618
Qy 721 SLVGTPLSAFCNLRPVLAEADQRLFTALPPEKNCNDNICQDDLSITFSPMSLDCLVVG 780
Db 619 TLVSEPLASFCNLRPVLAEADQRLFTALPPEKNCNDNICQDDLSITFSPMSLDCLVVG 678
Qy 781 GPRFNVTVVRNCGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWL-ACESASSTE 839
Db 679 GPQDFNMSVTLNCGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWL-ACESASSTE 738
Qy 840 VSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 899
Db 739 GHGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 798
Qy 900 FQLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHOVQVNLGORSIPISLVEVLP 959
Db 799 FQLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHOVQVNLGORSIPISLVEVLP 858
Qy 960 VRLNCTVTVDRPOVPSSENLSTCTKELPSHSDFLAELRKA PVNCSIAVCORIQCDI 1019
Db 859 VQINNVTVMDHPQVIFSNLSACHTEORSPPHSNFRDQRTVPVNCISVAVCKRIQCDL 918
Qy 1020 PFGIOEENATLKNLGFQVYKTSNHLIIVSTAELFNDSVFTLLPGQAFVRSQTE 1079
Db 919 PSFNTQVQVYVTKNLSFDVYKTSNHLIIVSTAELFNDSVFTLLPGQAFVRSQTE 978
Qy 1080 TKVEPFEVNPPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1137
Db 979 TKVEPFEVNPPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMWSEGGPPGAEPO 1036

RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
ID Q28984
AC Q28984;
DF 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactylia; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA SEQUENCE FROM N.A.
RP Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
```

```
DR GO:0008305; C:integrin complex; IEA.
DR GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00234; VWFPA; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 54.9%; Score 3815.5; DB 6; Length 920;
Best Local Similarity 79.4%; Pred. No. 1.4e-269; Indels 1; Gaps 1;
Matches 731; Conservative 80; Mismatches 109;

Qy 118 PQKFPALRGCPQBDSDIAFLIDSGSIIPHDPFRMKFVSTYMEQKKSKTLFSLAQYS 177
Db 1 PQKFPALRGCPQBDSDIAFLIDSGSIIPHDPFRMKFVSTYMEQKKSKTLFSLAQYS 60
Qy 178 EPRRIHTPKFPPONNPNRSLVKPIQLLGRTHATATGKVRBELNITNGAKNAFKIL 237
Db 61 EDPYTHFTFNDPKRNPSPKLLVRPIQLLGRTHATATGKVRBELFHSKSGARENAKIL 120
Qy 238 IVITDGEKFDPLGYEDVPEADREGVIRVVGDAFRSEKSRQSLNTIASKPPRDHPV 297
Db 121 VVITDGEKFDPLGYEDVPEADREGVIRVVGDAFNWSKSEELNTIASKPPGDHPV 180
Qy 298 QVNNFEALKTIONLREKIPIAEGTGTGSSSEHEHMSQEGFSAATISNGPLISVGSYD 357
Db 181 QVNNFEAVKTIQLOKTEPIAEGTGTGSSSEHEHMSQEGFSAATISNGPLISVGSYD 240
Qy 358 WAGVGLYTSKEKSTFTINMTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYQHIGLV 417
Db 241 WAGVGLYTSKEKSTFTINMTRVDSMDNDAYLGVAALILNRVQSLVGLAPRYQHIGLV 300
Qy 418 MFRONTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQV 477
Db 301 MFKQNSGAMKNAIKGSIQSGVFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQV 360
Qy 478 CPLPRGORARWQCDVAVLYGQGPWGRFGAALTVLGVDVNGKLTVDVAIGAPGBEDNRG 537
Db 361 CPLPQG-RAKWQCRVILCGEGHPWSPRFGAALTALGVDVNGKLTVDVAIGAPGBEDNRG 419
Qy 538 YLFHTSGSGISPSHSORLAKSLSPRLQYFGOSLGGQDITMDGLVDLTVGAQGHVLL 597
Db 420 YLFHTSGSGISPSHSORLAKSLSPRLQYFGOSLGGQDITMDGLVDLTVGAQGHVLL 479
Qy 598 RSQVPLRVKAIMENPREVARNPECDQVVKKGAGEVRVCLVQKSTRDRLAEGIOQS 657
Db 480 RSQVPLRVKAIMENPREVARNPECDQVVKKGAGEVRVCLVQKSTRDRLAEGIOQS 539
Qy 658 VVITYDALDGRPHSRVAFNETKSTRTOVGLTQTCETLKQLPNCIEDPVSPIVL 717
Db 540 VVITYDALDGRPHSRVAFNETKSTRTOVGLTQTCETLKQLPNCIEDPVSPIVL 599
Qy 718 LNFSLVGTPLSAFCNLRPVLAEADQRLFTALPPEKNCNDNICQDDLSITFSPMSLDC 777
Db 600 LNFSLVGTPLSAFCNLRPVLAEADQRLFTALPPEKNCNDNICQDDLSITFSPMSLDC 659
Qy 778 VVGPRFENVTVVRNCGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESASS 837
Db 660 VVGPRFENVTVVRNCGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESASS 719
Qy 838 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
Db 720 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNK 779
Qy 898 TEFQLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHOVQVNLGORSIPISLVEL 957
```

```

780 TEFQLELPVKAVVWVTSLEVSTKYENFTASEKTRHVIHQVFNGLGQKLPISVVFV 839
958 VPVRLNQTIVNDPQVTFSENLSTCTKRLPSHSDFLAELRKAPVYVNCISAVCORIQ 1017
840 VPVRLNRTVWDQVQVTFQNLSSCTEELGPRHSDFLKLOKTPVNLNCISAVCQKIQ 899
1018 DIPFGIQEENATLKGNLSP 1038
900 DIPSGIQEELKVTILKGNLSP 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
ID Q81VA6;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; 1A7B484FEFC79EB6 CRC64;

```

Query Match 59.3%; Score 3484; DB 4; Length 1169;
 Best local Similarity 61.5%; Pred. No. 3 5e-245;
 Matches 694; Conservative 138; Mismatches 291; Indels 6; Gaps 4;

```

QY 1 FNLDTENAMTFQENARFGQSVVQLGSRVYVVGAPQEIIVANQSGLYQCDYSTGSCPI 60
20 FNLDTEELTAFRVDSAGFGDSVVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACPEI 79
61 RLOVPVEANVMSLGLSLAATTSPQOLLACGPTVHQCSENTYVKGICPLFGSLNLRQPOK 120
80 GLOVPEANVMSLGLSLASTSPQOLLACGPTVHCEGGRNMTLTGLCLGPT -QLTOR 137
121 FPEALRGCPQEOESDIAFLIDGSGIIPDPRMKEFVSTVMEQLKSKTLPSLMQVSEBF 180
138 LPSVRECPQEOQDIIVFLIDGSGSISSRNFAVNFVRAVISQFQRPSTQFSLMQPSNKF 197
181 RIHTFKEFQNNPNPSLVKPTOLLGRTHATGVRKVIRELLNITNGARKAKFLIYI 240
198 QTHFTFEFRSRNPISLLASVHQLGQFTYTAIGNVVRHLPFASYGARRAAKFLIYI 257
241 TDCEKPGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNITASKPRDHVQYN 300
258 TDCKEGBSDLYKDVLPFADMAAGIIRYAGVGLAFQNRNSWKNELNDIAKSPQSEHIFKVE 317
301 NFPAKTIQNLQEKIPIAEGTGTGSSSPHEHMSQEGSAITSGPLISTVGSVDNAG 360
318 DFDALQIQNLQEKIPIAEGTGTGSSSPHEHMSQEGSAITSGPLISTVGSVDNAG 377

```

```

QY 361 GVFLYTSKEKSTPIKTRVDSMDNAYLGYAAATILNRVQSIAMVLGAPYQHIGLVAMER 420
378 GAFLYPPNNMSPFIINMSQENVDMSYLGYSYELALWKGVQSLVLCAPRYQHTKAVIF 437
421 QNTGMWESNANVKGTOIGAYFGASCLSDVDVDSNGSTDVLVLI GAPHYTYBTGCGQSVCP 480
438 QVSRQWRMEAEVTGTOIGSYFGASCLSDVDVDSNGSTDVLVLI GAPHYTYBTGCGQSVCP 497
481 PRQORARWOCDAVLYGEOGQPHGRFGAALTULGDVNGDKLTDVAIGAPCEENRGAIVYLF 540
498 PRGWR - RWWCDVLYGEOGHPWRFGAALTULGDVNGDKLTDVVI GAPEKEENRGAIVYLF 556
541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVLT VTAQGQHVLLRSQ 600
557 HGVLPFSISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVLT VTAQGQHVLLRSQ 616
601 PVLVRKALMEFNPREVARNVPECDVYVKGKEAGEVVRVCLHVOKS TRDLRBEQIQSVVT 660
617 PVLWVGVSQFIPAEIPRSAFECREQVSEQTIVQSNICLYIDKRKNLGLSGRDLQSSVT 676
661 YDLALDGRPHSRAVFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
677 LDIALDPGRSLPRATQETKNSRSLSRVVLGLKHCENFNLLPSCVEDSVTPITLRNLF 736
721 SLVGTPLSAFNGNLRPVLAEDAQELPTALPFFKNCNDNICODDL SITPSFMSLDCLVVG 780
737 TLVGKPLLAFLNLRPMLAADAQRYFTASLPFFKNCADHICODN LGISFPFGLKSLVG 796
781 GPREFNVTVVRNDGDSYRTQVTEFFPLDLSVRKYSTLQNRQSQRSLACSSASSTEV 840
797 SNLELNENVWVNDGDSYRTQVTEFFPLDLSVRKYSTLQNRQSQRSLACSSASSTEV 854
841 SGALKSTSCSINHPPIPPENSEVNTPTFDVDSKASIGNKLLKANTVSENNMPTNKTEF 900
855 SQGTWSTSCRINHLIPFGGAQITFLATFDVSPKAVLGDRLLLTANVSENNMPTNKTEF 914
901 QLELPVKAVYVWVTSLEVSTKYENFTASEKTRHVIHQVFNGLGQKLPISVVFV 959
915 QLELPVKAVYVWVTSLEVSTKYENFTASEKTRHVIHQVFNGLGQKLPISVVFV 974
960 VRLNQTIVNDPQVTFSENLSTCTKRLPSHSDFLAELRKAPVYVNCISAVCORIQCDI 1019
975 VELNQEAVMVDVESHVHPQNPQLSRCSSEKIAPPASDFLAHIQKNPVLDSCSIAGCLRFCDV 1034
1020 PFGIQEENATLKGNLSPDYIKTSNHLNLLIYSTAEILENDSVFTLLPGQAGAFVRSQTE 1079
1035 PPSVQEEELDTLKGNLSPFGVRQILQKTVSVWSVAEITFDTSVYSQLPQGEAFWRAQTT 1094
1080 TKVPEPEVNPPLIVGSSVGGILLALLALITAAALYKLGFFKQYKDMSE 1128
1095 TVLEKVKVHNFTPLIVGSSVGGILLALLALITAAALYKLGFFKQYKDMSE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
ID Q9QXH4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Teeng S.-Y., Pardoll D.,
RA Tsuchiya H.;
RT Isolation of Genes Selectively Expressed by Dendritic Cells.11;
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AF211864; AAF23492.1; -.

```


DR HSP; P11215; 1BHQ.
DR MGD; MGI:96609; ItgaX.
DR GO; GO:0008305; C: integrin complex; IEA.
DR GO; GO:0004835; F: cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P: cell-matrix adhesion; IEA.
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VVFA; 1.
DR Integrin.
KW SEQUENCE
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;
Query Match 56.6%; Score 3326.5; DB 11; Length 1169;
Best Local Similarity 57.2%; Pred. No. 1.1e-233;
Matches 652; Conservative 167; Mismatches 302; Indels 19; Gaps 7;
Qy 1 FNLDENATFTQENARGCGVQVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPEI 60
Db 20 FNLDAAELKTHFMDGAEFGHSLVQLQDSRWVVGAPKEIKATNQIGGLYKGYHTGCEPI 79
Qy 61 RLQVEEAVNMSLGLSLAATTSPPQLACGPTVHTQTSNTYVKGCLFLFGSNLRQQPQK 120
Db 80 SLQVPPPEAVNISLGLSLAATNPSSWLLACGPTVHTCRENLYLTGLCLFLSSSPKQ-QN 138
Qy 121 PSEALRGCPQSDIAFLIDGSGIILPHDPRMEKFYSTVNEQLKSKTLFSLMQVSEEF 180
Db 139 PETAQCECPQDDIVFLIDGSGISSTDFEKLDFVAVVMSQLQRESTRFSLMQFSDYF 198
Qy 181 RIHFTFEFQNNPRSLVAKETITQLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 199 RVHFTFNNFISTSPSLSLGSRVQLRGVYTTASAKHIVITELFTTQSGARQATKVLIVI 258
Qy 241 TDGEKFGDPLGYEDVIPAEDREGVIRVYVIGDAPRSEKSEQLNTIASKPRDHVQVN 300
Db 259 TDRKQGNLSYDSVIPMAEASIRIYALGVKAPYNEHSKQELKATASMPSEHVSVE 318
Qy 301 NFEALKTQNLREKIFAETGTQSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAQ 360
Db 319 NFDALKDIENQLKEKIFAETETPSSSTFELEMSQEGFSAVTPDGPVLGAVGSPWSG 378
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLYGAAAILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 379 GAFLYPSNKRPTFINNSQENEDMRDAYLGYSTALAFWKVGHSLILGAPRHQHTGKVIET 438
Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLVLGAPHYEYBTRGGQSVSCPL 480
Db 439 QESRHWREPKSEVKGTOIGSYFGASLCSDVDNDSGSTDVLVLGVPHYEYBTRGGQSVSCFM 498
Qy 481 PRGQARQCDVAVLYGEOQPMWREGAALTVLGVNGDKLTDVAIGAPGEEDNMGAVLYF 540
Db 499 P-GVGRWHGCGTTLHGEQCHPWRFGAALTVLGVNGDPSLADVAIGAPGEENRGAVYF 557
Qy 541 HCTSSGSISSPSHQRIACKSLSPRLOYGOSLGGQDILTMGLVNDLTVGAQGHVLLASQ 600
Db 558 HGASRQDAPSPSQISASIPSRIOYFGQSLGGQDILTRDGLVDLAVGSRVLLKTR 617
Qy 601 PVLIRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRSGQTSVYT 660
Db 618 PILRVSPVTHFTPAIRLSKSVFCQSQVAPQETLSDATVCLHIESPKTQL--GDLSRTV 675
Qy 661 YDLALDSGRPHRAVFNETKSTRQTVGLTQTCETLKQLPNCIEDPSPVIRLNF 720
Db 676 FDLALDHGRSLSTRAIFKETKTRALTRVKTGLGNHCESYKLLLPACVEDSVPTILRNF 735
Qy 721 SLVGTPLSAFGNLRPVLAEDAQLFTALFPFEKNCNDNICQDDLSITFFSMSLDCLVVG 780

Db 736 SLVGVPISSLOQLQPLAVDDQTYFTASLPFEKNCADHICQDDLSVVGFPDLKTLVVG 795
Qy 781 GREFNVTVTVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQ-----NORSQSWR 829
Db 796 SDLELVNDVTVSNDGSDSYGTTVTLFPVGLSPRVAEGQVFLRKEDQOWQRGQSHLH 855
Qy 830 LACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITEDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLMSTSCSRHVIIFRGSGSQMTFLVTFDVSFPAELGKLLLRARVGE 913
Qy 890 NNMPTNKTEFOLELPVKYAVVAVVTVSHGVSTKYLNFATSE-NTSRVMQHYQVSNLQOR 948
Db 914 NNVPCTPKTTFQLELPVKYAVVTVMISSHDQFTKYLNFSTSEKETSVEHRRFQVNNLQOR 973
Qy 949 SLDISLVFLVPLNQTIVDRPQVTFSENLSTCTKELPSHSDFLASLRKAPVNVCS 1008
Db 974 DVEFSINFVPIELKEAVN-TVMVSHQPQPLTQCYRNLKPTQFDLLTDMOKSPVLDGS 1032
Qy 1009 IAVCQRIQCDIPFGIOEBFNATLKNLGSFDWKYKTSNHNHLLIVSTABILFNDSVFTLLP 1068
Db 1033 IADCLHRCDDIPSLGILDELIFILKGNLSPGWTISQTLQKVVLLSEBITFTSVVSQLP 1092
Qy 1069 GQAFVRSQETETKVEPFEPVNPPLPLIVGSSVGLLALLAITAALYKLGPPKROYKDMSE 1128
Db 1093 GQAFILRAQTKTVLEMYKVHNPVPLIVGSSVGLLALLAITAILYKAGFPKROYKEMLEE 1152
RESULT 6
QSOYE7 PRELIMINARY; PRT; 1161 AA.
AC QSOYE7;
DT 03-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., Vandervieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -
DR HSP; P11215; 1BHQ.
DR GO; GO:0008305; C: integrin complex; IEA.
DR GO; GO:0004895; F: cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P: cell-matrix adhesion; IEA.
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VVFA; 1.
DR Integrin.
KW SEQUENCE
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;
Query Match 55.3%; Score 3251.5; DB 11; Length 1161;
Best Local Similarity 58.0%; Pred. No. 3.4e-228;
Matches 654; Conservative 158; Mismatches 302; Indels 13; Gaps 9;
Qy 2 NLDTENATFTQENARGCGVQVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPEI 61

Db 21 NLDVEPIVREDAAFGQTVVQFGSSRLVVGAPLEAVAVNQTGRLYDCAPATGMCPQIV 80
Qy 62 LOVPVEAVNMSLGLSLAAATTSPPQLLACGPTVHQTCSNTYVKGCLFGLSGNLROPOKPF 121
Db 81 LRSPLAVNMSLGLSLVATNAQLLACGPTAQACVKNYAKGSCLLGLGSL-OFIQAV 139
Qy 122 PRALRCGPQEDSDIAFLIDGSGSIIPHDPRRMBKFVSTVNEOLKSKTKLFLSLMOYSEFR 181
Db 140 PASMPCPRQEDMDIAFLIDGSGSINORDPAQMKDFVKALMGFASTSTLFLSLMOYSEFR 199
Qy 182 ILETFKBFONNPNRSLVLPDITOLLGRTHTATGVRKVIKRELLNITNGARKNAFKILIVIT 241
Db 200 THFTPTFEKNIIDPQSLVDPIVOLQGLYTATGIRTWMBELFHSKXGSKSAKILLIVIT 259
Qy 242 DGEKFDPLGDEVDIPEADREGVIRYVIGVDAFSEKSKRQBLNTIASKPPRHHVQVNN 301
Db 260 DQKYRDLPEYSDVIPAADKAGIIRVAIGVDAFQEPALKEINTIGSAPPQDRHVKN 319
Qy 302 PEALNTIONLREKIPALBGTQSGSSSPHEHMSQEGPSAATTSNGPLLSITVGSYDWDAG 361
Db 320 FAALRSIQOLQKFIKFAIBGTQSRSSSPHEHMSQEGPSAATTSNGPLLSITVGSYDWDAG 379
Qy 362 VELYTSKEKSTFTNTRVDSMDMDAYLGYAAAIIILNRRVQSLVGLGAPRYQHIGLVAMFRQ 421
Db 380 AELYPNTPTFTINMSQENVDMEDSYLGYSTAVAFWKGVHSLILGAPRHOHTGKVIFPQ 439
Qy 422 NTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYKQREGGOVSCPLP 481
Db 440 EARHWRPKSEVRGTQISGYFGASLCSVDVDRDGTDLVLIGAPHYKQREGGOVSCPLP 499
Qy 482 RGRARWQCDVLYGQCGQWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNRGAVLYFH 541
Db 500 -GVRGWQCEATLHGQHPWGEFVALTVLGDVNGDNLADVAIGAPGEESRGAVLYFH 558
Qy 542 GTSGSGISPSHQRISAGLSPLQVFGQSLSGQDLTMDGLVLTIVGAGHVLRLRSP 601
Db 559 GASRLBIMPSRQVGTGSLSLRQVFGQSLSGQDLTMDGLVLTIVGAGHVLRLRSP 618
Qy 602 VLRVKAIMEFNPBEVARNVPECDQVGVKGEAGEVVRVCLHVOKSTRDLREGQISVVTY 661
Db 619 LLKVELSIRPAPNEVAKAVYQWERTPTVLEAGEAVCLTVHKGSPDLL--GNVQGSVRY 676
Qy 662 DLALDGRPHSRVAFNETKXSTRQVLTGLTQTCETLKLQIPNCIEDVPSVIVRLNFS 721
Db 677 DLALDGRLLISRAIFDETCKTLGRLGHPCHETVKLLPDKCEDVAVSPILNFS 736
Qy 722 LACTPLSAFNLPLAEDAQRLPTALPFEKNCNDNLCODDLSTPFSMSLDCIACVCG 781
Db 737 LVREDSAP--RNLPVLAVGSDHITASLPEKNCQKELLCEGLGSLFSGLOVLVWG 795
Qy 782 PREFNVTVTVRNDGSDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACASASTEVS 841
Db 796 SPELTVTVTVWNEGSDSYGLVKFYFAGLSYRVVTGTQ-QPHQYPLRLACEAPAAQED 854
Qy 842 GALKSTSCSINHPIPENSEVTNIPEDVDSKASLGNKLLKANVTSENMBRTNKTETQ 901
Db 855 --LESSSCSINHPIFREGAATITMTFDVSYKAFGLDRLLLRKASSENKNDPTNKTAFQ 912
Qy 902 LELPKVAVVWVTSHGVSKYILNFASENTSR--VMQHOYQVNLGQSLPISLVLVFP 960
Db 913 LELPKVAVVWVTSHGVSKYILNFASENTSR--VMQHOYQVNLGQSLPISLVLVFP 972
Qy 961 RLNQTVIWRDPQVTFENLSS--TCHTKERLPSHSDLAELRKPAPVNCVIAVQRCQDI 1019
Db 973 LLNGVAVWD--VTLSSPAQGVSCVQKMPQPPDFLTQIRRSVLDSCDIADCLHFRCDI 1029
Qy 1020 PFFGIQEFENATLKNLSPDWYIKTSHNHLIIVSTABILLFNDVSVFTLLPQOGAFVRSQTE 1079
Db 1030 PSDIQDELFLRGNLSPFWVQVQTLQEKVLLVSEAEITPDTSVYQLPQGEAFBAQVBE 1089
Qy 1080 TKVEPFPVNPPLIIVGSSVGGLLALITAAIYKLGFFKRYQKMM 1126
Db 1090 TTLEEVVYVEPIFLVAGSSVGGLLALITAAIYKLGFFKRYQKMM 1136

RESULT 7

QSWTV4
ID Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; --
DR HSP; P20701.1; IUPA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR PROSITE; PS0234; VWFA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B1027B209E431 CRC64;

Query Match 26.0%; Score 1529.5; DB 11; Length 1161;

Best Local Similarity 33.8%; Pred. No. 2.5e-102;

Matches 396; Conservative 220; Mismatches 456; Indels 101; Gaps 37;

Qy 1 FNLDTENAMTQENA-RFGQSVVQLQGSRVVVGAPQEIIVANQEGSLYQCDYSGSCBP 59
Db 24 YLNDTRPTQSLAQAGRHFGYQVLIQEDG-VVVGAPGE---GDNTGGYHCHTSSEFCQ 79
Qy 60 TELQVPEAVNMSLGLSLAAATTSPPQLLACGPTVHQTCSNTYVKGCLFGLSGNLROQ 119
Db 80 VSLH-GSNHETSKYLGXTLATDAKGSLLACDPLSRTCDQNTYLSGLCVLPQSLGPM 138
Qy 120 KPPEALRCGPQEDSDIAFLIDGSGSIIPHDPRRMBKFVSTVNEOLKSKTKLFLSLMOY 179
Db 139 QNRPAYQECMKGVLDLVELFDGSSQSLDRKFEKILEMKDVNRKLSNTSYQFAAVQFSD 198
Qy 180 FRIHPTFKEP-QNNPNRSLVLPDITOLLGRTHTATGVRKVIKRELLNITNGARKNAFKIL 238
Db 199 CRTEFTFDYIKQNKPNPVLGSGVQPMFLNTFPAIVVAVHVKESGAPPDATKVLV 258
Qy 239 VITDGEKFGDPLGDEVDIPEADREG-----VIRYVIGVDAFSEKSKRQBLNTIASK 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIICIGHFVSVQKQKTLHFASBP 304
Qy 292 PRDHVQVNNFEAKTIQNLREKIFAETGQTGSSSPHEHMSQEGPSAATTSNGPLLS 351
Db 305 VEFPKILDTFEKLDLFTDQRRIIAYIEGTRNQLDTFNMELSSSGISADLSKHAVUG 364
Qy 352 TVGSYDWDAGVFP-LYTSKEKSTFTNTRVDSMDMDAYLGYAAA-IIILNRRVQSLVGLGAP 409
Db 365 AVGAKWDAGGLDLREDIQQGATFVQCEPLTSDVRGGYLGTVVAMTSSRSSRPLLAGAPR 424
Qy 410 YOHIGLVAMFR--QNTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLIGAPHY 467

Db 775 ANLTSSPARS-----GPRLKSSASLAVETLNSGSDAYVVRDLDPFRGLSPRKV 827
Qy 817 STLQNRQORSWRLACESASTESVSGAL-KSTCSINHPPIPESENTEFNITDVEDSKAS 875
Db 828 ENLQ---PHSRMPVSECEL--TEGSLTKIKNKVSSPIKAGQVSLQVFMFLNLS 882
Qy 876 LGKLLKANVTSEN-NMPTNKTEFOLELPVKYAVMVVTSHGVSSTKYLNTASENISR 934
Db 883 WEDFVELNGTVHCENENSLQEDNSAATHIPVLPVNLTKQENSTLYISFTPKGPKTQ 942
Qy 935 VMOHOYQVSNLQORSLSPLSLVLPVRLNQTWIDRPO-----VTFSENLS--TCHTK 986
Db 943 QVORVYQV-----RIQSAVDNMT-LEALVGVPRPHSDLIITVMSVQDPLVTCHSE 996
Qy 987 E-RLPSSHDFLAELKAPVNCSTAVCORIQCDIPFFGIGQEEFNATKGNLSFDWYIKTS 1045
Db 997 DLKRP-----SEAPCLPGV--QFRCPV-----RWEILQVGTWLSKEIKAS 1041
Qy 1046 HNLHIVSTAEILLFNDSTVFTLLPGGAFVRSQTEKVPFVFNPLPLVGVSSVGLL 1105
Db 1042 -STLSCLSSVSFNSKHFHLYGSKA-SEAQVLVKVDLIHEKMLHVVYLSGGGLVL 1099
Qy 1106 ALITAAALYKGFPRQYKDMW-SBGGPQCAEP 1136
Db 1100 FLIFLALYKVGFFKRLKKEADGGVNGSP 1131

RESULT 9
Q98TF1 PRELIMINARY; PRT; 1196 AA.
AC Q98TF1
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE CD11-1.
DE CIAL.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_Taxid=7962;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048536; BAB39134.1; -.
DR HSSP; P20701; 1LFA.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Aspartic-type AS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 24.1%; Score 1414; DB 13; Length 1196;
Best Local Similarity 31.3%; Pred. No. 7,2e-94;
Matches 373; Conservative 227; Mismatches 446; Indels 144; Gaps 39;
Qy 1 FNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQETVAANQORGLYQCDYSTGS 56

Db 32 FNIDTEHLRFNGAPEDFFGYSVYQTEFGNKRQIIVGAPLE---GNSGTMYSTADLOS 88
Qy 57 CEPIRLOVP-----VEAVNMSLGLSLAATSPPLLACGPTVHOTCSBNTVYVGLCFLFGS 112
Db 89 CQ--RLQRPQGSVRPFQMSAAVSSAALTS-----CSPYFPECDSNVLNGVCYQFSS 140
Qy 113 NLRQOPKPFPEALRGCPQEDSDIAFLIDSGSIIPHDPRRMKGFVSTVMOLKKSKTLFS 172
Db 141 SL-QAVSNFTAAVQECSEKREVNLPFLFDGSSMKTVDFENKXNFIDIMKLSNLSKFA 199
Qy 173 LMQYSSEFRHPHTKFPQNNPNSLVKPIITOLLGRHTATGVKVIKRELLN-ITNGARK 231
Db 200 AVQSTDTVRVTFPNDVQSGSABEKLWKE--THKSLTNTHKADYILNKLNSKLSADS 258
Qy 232 NAYKILIVITDGKFGEDPLGYED--VPEADREGVIRYVIGVGDADFSEKSRKRELNTIAS 289
Db 259 KAQKALVITD---GDPSSNDDBYVNLKKCKDEQNLIRLYIIGV-----KVDLIETQLAS 309
Qy 290 KPRDRHVQVNNPEALKTIONQLREKIFATEGTQVGSSSFHEHXSQBGSAALTSNGPL 349
Db 310 EPRKNNTFYIKQYSGLGLLDNLQKLIYNEGSDVAGQDRQKELSGSGSVVYQBSVI 369
Qy 350 LSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMN-DAYLGIAAAIILNRVOSLVIGAP 408
Db 370 VGSVGSNDWEGALYEVTG--SGSDFKETIIDPAVNDKSYMGYSTVLGMRGVSLLFSGAP 428
Qy 409 RYOHIGLVAMFRONTGMBSNANVKGTOICAYFGASLCSDVDNSGSTDVLIGAPHYH 468
Db 429 RAEHTGLVTLFTKNQNTWVTSNINGEQIGSYFGASLSLDDVSDGSDLLVGAFLFYQ 488
Qy 469 QTRGGQVSVCPLRGQRRQARMCDAVLYGEQ-----GQPMGRFGAALTVLGDVNGDKLTDV 523
Db 489 SQ-----PRTEGLRVYVLSLSEQKYFKTLNVQSQTTCGRFAASVASLXDLNGDGLSDV 540
Qy 524 AIGAPGEDNRGAVYLFHGTSGSGISPSHS--QRIAGSKLSPRIQYFQCSLSGGDLTWDG 582
Db 541 AVGAPLE--NEGVVYIYLGDRTHGINPELTPQISVQSVLPGLCQFGVSTQGMNDN 598
Qy 583 LVDLTGCAQGHVLLLRSSQVPLRVKAIMEPNPREVARNVFECNDQVKGKEAGEVRVCLHV 642
Db 599 LTDIVGAGQGVILLKARFVMSYSAQLSPSPKISLNYFECPS--NAPNAFLNLTSCFTV 656
Qy 643 QKSTRDLREGQIQSV--VTVYDLALDSGRPHSRVAVNETKNSTR--RQTVGLGTQTCET 698
Db 657 TERTSS---TGSLEKKNVSLNLDVVRGMSRGFFPDQSSVSRSTLQOQSVLLDSSGSCFN 713
Qy 699 LKQLPNCIEDPVSVILRLNESILVGTPLSAFGLNLRPLAEDAQLFTALPPPEKNCND 758
Db 714 FSPMLRCVADTVSPKIRNFS--QTEMLS--GNSVAVLVDHSEKTEENVEVFFQRCN 770
Qy 759 NICODDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSYRTQVTPFPPLDLSYKXST 818
Db 771 NSCVADLKLNFSTN--NTLVVENQAHFTVQVSLANPGDSDSYNTSIVLHYPBGISLSKFA 829
Qy 819 LQNRORSWRLACESASTESVSGALSKSTCSINHPPIPESENTEFNITDVE---DSKAS 875
Db 830 IKPSRTR-----SSCGDRDSCGATRTTCSIDLPIVRSCTTQFGLGTFVGMKWDNDS 881
Qy 876 LGKLLKANVTSENMPNTRNKTEFOLELPVKYAVMVVTSHGVS-STKYLNTASENISR 934
Db 882 NEMEIMITANSNNGM---SDTEVRSVFPVQFAVDLAISLVAEDSVTYMNFSLDRGPK 938
Qy 935 VMOHOYQVSNLQORSLSPLSLVLPVRLNQTWIDRPO-----VTFSENLS--TCHTK 986
Db 939 PLANITYKVENSGKDLFVSVTLPLCQPTPHVLTPTFTSMHEVHHSFISSYHQIIMCLLN 998
Qy 964 QTVIWDPRQVTFSENLSSTCHTKERLPSHGDFLAELKAPVNCSTAVCORIQ-----C 1017
Db 999 KHLFFSFELSVAQVVRSTGWSLREVC-SQP--DLNKSSAVHPLTADARLQNVKEYES 1055
Qy 1018 DIPFGIGEE--FNATLKNLSFDWYIKTSNHLHIVSTAEILLVSTAILPNDVSTFILLPQCGAFVR 1075

Db	1056	KYSFTEFRKDVNSISAEINNTSLYNQTS-----SELKYNPH-----	R 1094
Qy	1076	SQTEKTVKVEFPVPLPLIV-GSSVGGLLHLLALITAAALYKLGPFKQYKD	1124
Db	1095	SQTEKTVKVE-FVVPPLMLIVCTGAVGGFFLIILITLLKRCGFFKRNRPD	1143
RESULT 10			
Q98TF0	PRELIMINARY; PRT; 1187 AA.		
AC	Q98TF0;		
DT	01-JUN-2001 (TRENBLrel. 17, Created)		
DT	01-JUN-2001 (TRENBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)		
DE	CD11-2.		
GN	C1A2.		
OS	Cyprinus carpio (Common carp).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Cyprinus.		
ON	NCBI_TaxID=7962;		
EX	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Peritoneal exudate cells;		
RA	Kimura M., Fujiki K., Nakao M.;		
RT	"Molecular cloning of a leukocyte integrin from the common carp.,"		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB048537; BAB39135.1; ..		
DR	HSSP; P20701; 1LFA.		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.		
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001369; Asparticase AS.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	InterPro; IPR002035; VWF A.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	Pfam; PF00357; integrin_A; 1.		
DR	Pfam; PF00092; vwa; 1.		
DR	PRINTS; PR01185; INTEGRINA.		
DR	PRINTS; PR00453; VWFADOMAIN.		
DR	SMART; SM00191; Int_alpha; 5.		
DR	SMART; SM00327; VWA; 1.		
DR	PROSITE; PS00141; ASP PROTEASE; 1.		
DR	PROSITE; PS0234; VWFA; 1.		
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;		
Query Watch			
Best local Similarity 31.5%; Score 1363.5; DB 13; Length 1187;			
Matches 375; Conservative 209; Mismatches 451; Indels 157; Gaps 43;			
Qy	1	FNLDTENAMTPOENARG-FGQSVVQLQ-GSR--VVVGAPQEIIVAANQRGSLYQCDYSTGS	56
Db	32	FNIDTEHLRFNGTPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSCCTADLQS	88
Qy	57	CEPIRLQVP-----VZAVNMSLGLSAATTSFPQLLAGCTVHQTCSNTYVKGCLFLGCS	112
Db	89	CK--RLQRFQSGSVRFFGMSAAVSAALTS-----CSYPFAHECDGNSLVNGVCYQFNS	140
Qy	113	NLRQCPQKPEALRGCPQSDIAFLDGSIIIPHDPRMKPFVSTWEOQLKSKTLFCS	172
Db	141	SL-QAVSNFTAAVQESKREVNLFVFDGSSMKAVEFDMKNFKIDVMKLSNSIIFKA	199
Qy	173	LMQYSEFRIHTFKFQONPNRPSIVKPIITOLLGHTTATGVRKVIKRLIN-ITNGARK	231
Db	200	AVQFSTEIRTFDFNDYQNSAEELMKR-RHMKSLNTYKAINTVLKNVLSVSGGADP	258
Qy	232	NAFKILVITDGEKFGDPLGYED--VIPADREGVIRYVIGVDAPFRSEKSRQELNTIAS	289
Db	259	NAQKALVIITD-----GDSNDNDYINILNCDQNILRYIIGV-----KVDLTTLTQLAA	309
Qy	290	KPPRDHVFQVNNFPAKTIQNLREKIPAIETGTQSGSSSFEHMSQSGFSAIISNGPL	349

RESULT 11
Q96HB1
ID Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
EX [1]

RP SEQUENCE FROM N.A.
RC Tissue=Lymph;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL, BCC08777; AA08777.1; --
DR GO: GO:0003305; C:integrin complex; IEA.
DR GO: GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO: GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWEA; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6FF2546B8C632F9 CRC64;

Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.2%; Pred. No. 2.7e-89;
Matches 376; Conservative 186; Mismatches 432; Indels 175; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARGFGSVVLOGSVVVGAPQIVANQSGLSYQCDYSTGSC 58
DB 26 YLNDVGASFPSPRAGRIFGRVQLV-GNGVIVGAPGE---GNSGSLYQCQSGTCHL 81
QY 59 PIRLQVPVAVNMISLGLSAAATSPOLLACGETVHQTCEHTYVYKGLCFPGSNLRQOP 118
DB 82 PVTLR-GSNVTSKYLGMTLA--TDP----- 103
QY 119 QKPEALRGCPQEDSDIAFLIDGSGSIIHPDFRMKEPVTVMQQLKSKTFLSLMOYSE 178
DB 104 -----TDGS-----ILFAVQVST 117
QY 179 EFRHFTFEFQNNPRSLVKEITQLGRTHATGVRKVIRELNITNGARKAPKILI 238
DB 118 SYKTEFDSYVRKDPDALLKHVKMLLTNTFGAINTVATEVFEELGARPDAFKVLI 177
QY 239 VIIDGKFGDPIGVYVPEADREGVIRVYVVGDAFRSEKSFQELNTIASPPRDHPVQ 298
DB 178 IITDGS--ATDSGNIIDAAND-----IIRYIIGKHQFQKESQETLHKFASKPASEFVKI 230
QY 299 VNMFEALKTQNLKRIKFAIEGTOTGSSSFHEMSQEGFSAATISNGELLSLVGSYDM 358
DB 231 LDTFEKLKOLFTELQKIVYIEGTSKQDLTSFNNELSSGISADLSRGHAVVGAVGAKDM 290
QY 359 AGGVF-LYTSKEKSTHINTRVDSMDNDAYLGAAA-IILNRVQSLVIGAPRYOHIGLV 416
DB 291 AGGFLDLKADLQDDTFIGNELTPEVRAGYLGVTVTLPSRQKTSLLASGAPRYQHNGRV 350
QY 417 AMFR--QNTGMWESNANVKQTQIGAYFGASLCSDVDNSGSTDLVLIGAPHYEOTRGQ 474
DB 351 LLLQEPQGGHNSQVQTHGTQIGSPFGELGVDVDQDETLLIGAPLFYGEQGR 410
QY 475 VSVCLPLRGQRARQCDAV--LYGEQGPQGRFGAALTULGVNNGKLTDLVAIGAGEED 532
DB 411 VFIV-----QRRQLGPEVSELOQDPCYPLGRFGEAITALTIDNGDLVDVAVGAPLEE- 464
QY 533 NRGAVLPHGTSGSGISPSHSIAGSKLSPRLQYFGQSLSGQDLMGLDVLDTVGAOG 592
DB 465 -QGVYIFNGHG-GLSFQSPQRIEGTVLSGQVWFGSRTHGVKOLEGDLADVAVGAES 522
QY 593 HVLLRSQPLVRKAIEMFNPREVARNVFCNDQV--KGKEAGEVRVCLHVQKSTDRRLR 651
DB 523 QMIVLSKRPVDMVTLMSPFAETPVHEVECSYSTSNKMEGVNITTCFQI-KSLIPQF- 580
QY 652 EGQIQSVTVYDLADSGRPHSAVFNKSTRQTOVLGLTQTCETLKLQLENCTEDPV 711
DB 581 QGRIVANITVTLQDXHRTARRRGLPGGRHRLRNIAVT-TSMSCDTDFSFHFPVCVDLI 639

QY 712 SPIVLRNFSI---VGTPLSAFGN-----LRPVLAEQAQRLFTALPPEKNCNNDNICO 762
DB 640 SPINVSINFSLWEEGTPRORACKDIPILRPSLHSETWEI-----PFEKNCCKEDKKE 694
QY 763 DDLSTTFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPFFLDLSYKRYSTLQNO 822
DB 695 ANLRVSPSPARSRALRLTAPASLSVLSLSNLEBDAYVQLDHFPPGLSFRKVEML--- 751
QY 823 RSQSRWLACHS--ASSTEVSGLKSTSCSINHPIPPENSESVTFNIPTDVSCKASLGNKL 880
DB 752 KPHSQIPVSCBELPEERLLSRAL---SCNVSSIFRAGHSVALQMMPTLVNSSWGSVD 808
QY 881 LKANVTSENN---MPRTNKTEPQLBLPVKIVYVMVTSHGVSSTKYLNFTASNTSRVM 936
DB 809 ELHANVTNNEDSDLEDNSATTI---IPILYPINILIQDOEDSTLYVSTFKPGPKIHQV 865
QY 937 QHQVQV---SNLQORSPL-ISLVELVPLNQTWINDRPQVTPPSENLSSTCHTK--BELP 990
DB 866 KMTQVRIQPSIHQHNITPLEAVVGPQPPSEGPITHQSVQMEPPV--PCHYEDELRP 923
QY 991 SHSD--FLAELRKAPVWNCSTAVCQRIQCDIPFPGIQEENATLKGNI-SFDWYIKTSHNH 1049
DB 924 DAARFCLPGALFRCPVV-----FRCEILVQVIGTLELVGEIEAS-SM 964
QY 1049 LLIVSTABILFNDVSFTLLPGQAFVRSOTKVEPPEVNPPLPLIVGSSVGGLLLLALI 1108
DB 965 FSLCSSLSISFSSSKHFLYGSNASL-AQVVMKVDVWYKQMLYLYLSGIGGGLLLLLL 1023
QY 1109 TAAVYKLGFFKQVKQKMMSEG-GPPGAP 1136
DB 1024 FIVLYKGVFFRNLKEMKAGRGVNGIP 1052

RESULT 12
Q8HZV0 PRELIMINARY; PRT; 927 AA.
AC Q8HZV0;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Sequence of the alpha subunit of bovine lymphocyte function-
associated antigen 1."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF440778; AAN63636.1; --
DR PIR; A32039; A32039.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWEA; 1.
FT NON_TER 1
FT NON_TER 927
SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match 21.6%; Score 1270; DB 6; Length 927;
Best Local Similarity 34.7%; Pred. No. 1.6e-83;
Matches 339; Conservative 167; Mismatches 386; Indels 86; Gaps 28;

Db 489 -EEDAFVRIEIEGOMSGYFSGVLCVPDIDMDGTTDFLLVAAPFYHIRGEGRVYVYQVE 547
QY 483 GQARWOCDAVLVYGGQWGRFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
Db 548 -ODASPSLAHTLSGPHLTNSRFGFAMAAGVDINDOKDKTDVAIGAPLSPFGAGDASVGS 606
QY 537 VYLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVGAQGHVLL 596
Db 607 VYIYNGHSG-GLYDPSQOIRASSVASGLHYFGMSVSGGLDFNGDGLADITVGRSDSAVV 665
QY 597 LRSQVPLRVKATMBENPREVARNVFCNDQVVKGEAGEVRVCLHVQKS---TRDLRREG 653
Db 666 LRSRPVDTLTVGNTTP-----DALPMWFIGM--DYNLCFVDSVVVASBGLREM 715
QY 654 QIOSVTVYDGLDSDGRPHSRAVENITKSTROTQVGLGTOTC----- 696
Db 716 FLNFTVDVDF-----TKQRRLQCEDSSGCSCLKRWGSGSLCEHFWLI 760
QY 697 EFLKQLPNCIEDPVSPIVRLNFSINGTPLSAPGNLR-----PVLAEQAQLFTALP--P 750
Db 761 STEEL-----CEEDCFSNITIKVYE-----POTSGRNDYFNPUL--DHYKEPSAIFOLP 809
QY 751 FFKNGCNDNICQDDLSITPSMSLDBLVGGGPRFNVTVVTRNDGEDSYRTQVTFPPFLD 810
Db 810 YEKDCNKVFCIAEIQLTN--ISQBELVVGVTKEVTWMTSLTNSGDSYTNMNLNYPEN 868
QY 811 LSYRKVSTLONORSQBSWPLACESASSTEVSALZSCSINHPIEPENSEVTENITFDV 870
Db 869 LOFKKI-----QKPSVPDQDDPKPV---ASVLVWNCIKIGHPIJ--KXSSVNVSVTQCL 918
QY 871 DSKASLGNKLLKANVTSNNPRTKTEFQLELPKAVVMVTVSHGVSTKYLNTASE 930
Db 919 ESSVFNRTADITVTSNNSKSLARETR--SLQPRHAFIAVLSR--PSVMYMN--TSQ 971
QY 931 NTSRVWQHOYVSNLQORSPLISLVLVPLVRLNQTIVWDRPQVTSNLSST-----CHT 985
Db 972 SPSDHKEPFFVHGENLFGAVFQLCVPIKLODF-----QIVRVKNLTKQDHTCTQ 1025
QY 986 KEPLSHSDFLAELKAPVNGSIAVCQRIQCDIPFGIOEFNATKGNLSFDWVTKTS 1045
Db 1026 SOEPACGSPVQVHXHWSVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILFNDSVFTLLPGQAFVRSQETKVBP-----EVPNPLPLIV 1095
Db 1066 HTKQLLRDVSLEPILGEISFNKSLYGLNAE-----NERTKITVIFLKEETRSPLII 1119
QY 1096 GSSVGLLILALITLALYKLGFFPKROYKDMSE 1128
Db 1120 GSSIGGLVLVLIILFKCGFFPKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Integrin alpha E2 (Fragment)
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between"
RT intraepithelial lymphocytes and dendritic cells in lymph."
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.

GO: GO:0008305; C:integrin complex; IEA.
GO: GO:0004895; P:cell adhesion receptor activity; IEA.
GO: GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01819; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
FIT NONTER 1
SQ SEQUENCE 1167 AA; 128593 MW; E3PED7E18B5CBBC CRC64;
Query Match 19.0%; Score 1119; DB 11; Length 1167;
Best Local Similarity 27.8%; Pred. No. 2.5e-72;
Matches 335; Conservative 204; Mismatches 423; Indels 244; Gaps 37;
QY 47 LYQCDYSTGCEPIRLQVPVEAVNNSLGLSAATSPQLLAGCPTVH----- 94
Db 67 LYQC-----ATSIIPDEIGCOFVEHILMPKRYQVGT 98
QY 95 -----QTCSENTYVRKLCPLFGSNLRQPKPEALRG----- 127
Db 99 LVNRHNGVLVCIQVSRKPSRLNSELATGACSLTLPNLDLQAAQYFSDLEGVLDLGASVNS 158
QY 128 -----CPQED-----SDIAPLIDGSGSIIPDPRMKFV 157
Db 159 GDYYSKGGSTGEBTKSARLRROAVEEEDERAGTEIAIVLDGSGSTEPSDFOKADFI 218
QY 158 STVMEQL--KKSCTLFSIMQYSEBFRIHFTPKFQNNPNPSLRVKPITQLLGRTHATGV 215
Db 219 STMEKVFKECFECNFALVQGVQIQTFFDLDSRDINASLAKVQSIQVQKVKITASAM 278
QY 216 RKVIRELLNIITNGARKNAFKLIVITGEEKGDDPGIEDVTPRADREGVIRVYVIGVDAP 275
Db 279 QHVLNIPISGRSKKALKVMVVLTDGIDIFRDLNLTIVISSKMQGVVFPALGVGNAP 338
QY 276 RSEKSRQELNTIASKPPREDHVQVNNFENLKTIONQIREKIFAETGTOTGSSSSPHEHMS 335
Db 339 ENNTYRELKLIADPKAAHFTKNTYNSALDGLSKLQORIIHMEGT---VGDITLYQLA 395
QY 336 QEGPSAAITSNGP-LLSTVGSYDWAAGGVFLY-TSKBKSFTFNMTVRVDSMDNA---YLGY 390
Db 396 QTGFSAQILDKGQVLLGTVGAFNWSGGALLYNTONGRGRFLNQT-AKEDFPRAAQYSYLYG 454
QY 391 AAAIILNRVOSLVLAGPYQHIGIVAMFRONTGMESNANVKGTQIGAYFGASLCSVDV 450
Db 455 SVALHKAHGVSYVAGAPRHKLRGAVFELQEDGEETFWRIEGEQMGSYFGSVLCPVDI 514
QY 451 DMSGTDVLVLCAPHYTEQTRGGQVSVCLPRGQARWQCDAVLYEGQGPWGRGAALT 510
Db 515 NWDGITDFLLVAAPFYHIRGEGRVYVYRV-HEQDAPPSLVTLTSGYPLGTSRPFAMA 573
QY 511 VLGDVNGDKLTDVAIGAP-----GEEDNRGANVLFHGTSGSGISPSHSQRIAGSKLSPR 564
Db 574 AVGDINDQDKFTDVAIGAFLEGFAGDAGSGSYGVYIYNGHSG-GLHASTSQIRASSVALG 632
QY 565 LQYFGQSLSGGQDLTMDGLVDLTGVGAQGHVLLLSRQVPLRVKATMBENPREVARNVFCN 624
Db 633 LYVFGMSVSGGLDFSGDLDLADITVGSQDVAVVLRSPVVDLTVMSTFTPDALP----- 685
QY 625 DQVVGKAGEVRVCLHVQKS---TRDLRREGQIQSVTVYDGLDLSGRPHSRAVENITKN 681
Db 686 ---NAFKDKMDYELCFKVDSSAIVPSEPLGRGLNSLNTFVDVDF-----TKQ 727
QY 682 STTRQ-----TQVGLGTQTCETLKLQLPNCIEDPVSPIVRLNAPS 721
Db 728 KQRLQCADRSQCSCLMKWSGSSSLCEHFLISTEEL-----CEDDFCSNITIKVSVE 781

QY	722	LVGTPLSAFNGRLPVLAEQAORLFTALF--PFKKCGNDNICQDDLSITFSFMSLCLVY	779
Db	782	F-QTSERRNHPIL--DHVKEPSAIFQLPYEKDCNKKVFCIAEIQUTTAISQOD-LVV	837
QY	780	GGPREPNVTVVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTE	839
Db	838	GIYKEVTNLSLNSGSDSYTNWALNYPNQLPKKI-----QKPLSPDIQDDPKPV--	890
QY	840	VSGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKE	899
Db	891	--ASVLWNCKIGHPII--KQSSVNVSVTQLEESIFFPNRTADITVTITSNNEKSLARETH	947
QY	900	FOLELPVKYAVYMYVTSHGVSTKYLNTFASENTSRVMQHOVQVSNLQORSLPISLVFLVP	959
Db	948	---SLOPRHAFIAYLSR--PSVMYNN--TQSSSDHKEFFENVGHNHFGAVTQLQICVP	1000
QY	960	VRANQTVINDRPQVTFSENISST-CHTKERLPSSHDFLAELKAPVNVCSIAVCQRIOCD	1018
Db	1001	ITLRDL-----QIIRVKHLTKTOAHT-----ECTQSQEPTCGSDPVQNV--	1039
QY	1019	IPFPGIOEEFNATL-----KGNLSFDWYIKTSHNHLI--VSTAELI-----FNDSVFTL	1066
Db	1040	-----EHWSVICAITSNKENTVAAETSMGHTKQLLRDISLQILGHSFNKSLIEG	1092
QY	1067	LPGCGAFVRSQTEPKVEPF-----EVPNPLPIVGSVGGLLLLALITAAIYKLGFFPKQY	1122
Db	1093	LNAB-----NHRTKITVIFLKEKPHSLPLIIGSSIGGLLVVITAILFKCGFFPKKY	1146
QY	1123	KDMWSE	1128
Db	1147	KQLNLE	1152

RESULT 15

Q7TQC3 PRELIMINARY; PRT; 1189 AA.

AC Q7TQC3

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE All integrin.

DE ITGALL

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A.

RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg B., Gullberg D.;

RA "alb1 integrin is important for mesenchymal cell function: elimination of alb1 leads to dwarfism."

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

EN [2]

RP SEQUENCE FROM N.A.

RA Johanson W., Popova S.N.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY124460; AAMG2130.1; -

EN Integrin.

EN KW

SO SEQUENCE

EN 1188 AA; 133067 MW; 69H2BF20D75E4271 CRC64;

Query Match	17.9%;	Score	1052.5;	DB	11;	Length	1188;
Best Local Similarity	28.2%;	Pred. No.	1.9e-67;				
Matches	349;	Conservative	184;	Mismatches	497;	Indels	207;
Gaps	45;						

Qy	1	FNLDTENAMTFQ-ENARFGFGQSVWQ--LQGRS-VVVGAPQIIVAAORGSIVYQCDYSTGGS	56
D5	23	FNMDTNPRVIAGSAAPFGVTVQOHDLSGKKWLWVGAPMETNQHGTQKTDYVKCPVTVQGN	82
Qy	57	CEPRL-----QVPEAVNMISGLSLAATSPPOLACGPTHQTCSENTIVYVGLCLFLF	110
D5	83	CTKLNIGRWLTLSNYSERKNNRGLSLATNPDKNSFLACSPLWSHGCGSSYTTGMGSRV	142

Qy	111	GSNLROQPOKFPEARLGCPOEDSDIAFLIDGSGSIIPHDFRRKNFVSTVMEQ--LKSK	168
Qy	112	QY	
Db	143	NSNFPSKTVAP-ALQRC-QTYMDIIVILVDGNSIYP--WVEYHFLINKPKYIPGPQ	198
Qy	169	TLFSLMOYSEBFIHFTFKEFONNPNRSLVXPITQLLGRTHATAGVRKVIPELLM	227
Db	199	IQVGIVQVEDAVHEPLNDYRSKDVENAASHIEQGGTEITATGIBFARSEAFQ--K	256
Qy	228	GARONAPXLIIVITGEKFGPLGYEDVITPEADREGVIRVICV-----GDARSEKSRQ	282
Db	257	GGKGAKKYMIIVITGESHDSP--DLEKVIQOGEKONVTRYAVAVLGYNNRRGINPETFLN	315
Qy	283	EANTIASKEPRDHVEOVNNFBEALKTIONLRBEKIPAIETGOTGSSSFHEHMQEFSAA	342
Db	316	EIKYIASDDPDHFFNVNTEAALKDVIDALGRIFSLGNTK--NETSPGLEMSOTGFSH	374
Qy	343	ITSNGELLSTVGSYDWAGVFLYTSKKB-----STFINMTRVDSMDNDVILGYAAAILR	397
Db	375	VVEDGILLGAVGAYDNAGVAKBETSAGKVIPIHRESYLKEPPEELKRVAVLGYTVTSWS	434
Qy	398	NRV-QSLVLGAPRYQHIGIVAMP-RONTGWESNANVKGTOICAYFGASCUSVDVDSNGS	455
Db	435	SROGRVYVAGAPRFNHTGKVIIFSMNNRSLTHQALRGEQIGSYFGSEBTSVDVNDRV	494
Qy	456	TDILVLCAPHYEQTR--GGQVSCPILPRGORAWOODAVLYGBQGPWGSFGAALTVLGD	514
Db	495	TJVLVVGAFWYSEGRGKVVYNL---RQNRFTVNGTLKDSHS-QNARFSGSIASVQD	551
Qy	515	VNGDKLITVAIGAPCEBDRGAVLYFHGTSGSGISPHSHORISAGSKLSPLQVFGQSLSG	574
Db	552	LQODSYNDVWVGAPLEDSHRGAIYIIFHGFO--TNILKKPMORITASELAPQLOHPGCSIH	610
Qy	575	QODLTMDGLDVLTVGAQGHVLLRSOPVARVKAIEMFNPREVARNVF--BCNDQWVKRGE	632
Db	611	QLDNEDELVDLAVGALGNNAVVLWARPVQVINASHFEPFSKI--NIFHKDC-----KR	661
Qy	633	AGEVRVCL-----HVOKSTRDLRSGOIQSVVTYDLDLDSGRPHSPAVFNET	679
Db	662	NERDATCLAAFLCFIPILAPHFQATVG-----IRYNATMDERYMPRAHLDGE	711
Qy	680	XNS--TRQTOVLGLTOTCETLKLQFNCEIDPVSPIVLRLNFSLVGTPLSAFGLMLRPVLA	738
Db	712	GDQFTNRAVLSSGQEQHCQRINEHVL-DADYKVPVAFSVEYSLEDP-----DNGFMLD	764
Qy	739	EDAQRULTALFPPEKXGNDNIQDDI-----SITFSF	771
Db	765	NGWPTTLRVSVPFWNGCNEDEHCVPDLVDARSDLPTAMEYCYQVLGRPAQDCSSYTLSP	824
Qy	772	MSLDCLVGSGPREFNVTVVRNDGOSYRQVTFPPPLDLSYRKVSTLQNRQORSKRLA	831
Db	825	DTTVFIESTRRVAVATLENGENAYSAVLMISQENLOF--ASLIQKDDSDNS--IE	880
Qy	832	CEASASTVSGALKSTCSINHPHIPPENSEVFNITFDVDSKASLEN--KLULLKANVTSEN	890
Db	881	CVNEER-----RLHKVCNVSVYPPRAKAKAVAPRLDPEFSKVFLHHLQHLGAGSDSHE	935
Qy	891	NMPRTNKTESQLBPLVKYIAVTWVTSHGUSTKYLNFTASENTS-----RVWQHYQV	942
Db	936	QDSTADNTALLAFHLKYEADVLFTFR---SSLSHFVEKANSLSIESYDGGPPNCVFKV	992
Qy	943	SNLQORSLPISLVL---VPV-----RL-----NOTVIMDRPQVT	974
Db	993	QNLG---FFPHGVNMKTIIVPIATRGNNELLVLRDPPTDQGNWSCNIGNSTEYRSTPI--	1048
Qy	975	FSENLSTCHTKRLPSHD---FLABLRKAPVWNCISAVCQRTQCDIPPGIGIEBFNAT	1031
Db	1049	-BEDLS---HAPQRNHSNDVWSIICNLRAP-----SQBTSFY	108
Qy	1032	LKGNLSDFYIKTSHNHLIVSTAEI---LFNDSVFETLLPQAGAFVRSQTKYKVPPEVP	108
Db	1084	LVGNLWJTSKALKIRSLKTVNALQORHSPFIF-----REEDPSRQVTFEIS	113
Qy	1089	N-----PFLIVGSSVGGLLLLALITAAALYKLGFFK	1119

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds
(without alignments)
6495.175 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKMMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 2823547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	2 AAW65090	AAW65090 Human Bet
2	5868	99.9	1153	3 AAB07360	AAB07360 Human CDI
3	5868	99.9	1153	5 AAU80252	AAU80252 Human int
4	5868	99.9	1153	5 AEG61469	ABG61469 Human Bet
5	5868	99.9	1153	5 AAO14428	AAO14428 Integrin
6	5868	99.9	1153	7 ADD25615	Add25615 Binding d
7	5858	99.7	1153	2 AAR04136	AAR04136 Alpha sub
8	3464	59.0	1153	2 AAR07120	AAR07120 p150.95 a
9	3450	58.7	1163	3 AAW65091	AAW65091 Human CDI
10	3450	58.7	1163	3 AAB07361	AAB07361 Human CDI
11	3450	58.7	1163	5 AEG61470	ABG61470 Human Bet
12	3448	58.7	1163	6 ABOU7406	ABOU7406 Protein d
13	3417	58.2	1161	2 AAR78166	AAR78166 Human bet
14	3417	58.2	1161	2 AAW23049	AAW23049 Human bet
15	3417	58.2	1161	2 AAW57491	AAW57491 Human bet
16	3417	58.2	1161	2 AAW5089	AAW5089 Human Bet
17	3417	58.2	1161	2 AAW72825	AAW72825 Human alp
18	3417	58.2	1161	2 AAW73342	AAW73342 Human alp
19	3417	58.2	1161	3 AAB07359	ABG61468 Human Bet
20	3417	58.2	1161	5 AEG61468	ABG61468 Human Bet
21	3401.5	57.9	1161	2 AAW23064	AAW23064 Human Bet
22	3401.5	57.9	1161	2 AAW5106	AAW5106 Human alp
23	3401.5	57.9	1161	2 AAW72837	AAW72837 Human alp
24	3401.5	57.9	1161	2 AAW73343	AAW73343 Human alp
25	3401.5	57.9	1161	3 AAB07376	AAB07376 Human alp

ALIGNMENTS

RESULT 1
AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX
AC AAW65090;
DT 28-SEP-1998 (first entry)
DE Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or

XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLGSSRVVVGAPQEIIVANQKSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLGSSRVVVGAPQEIIVANQKSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEANVMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSLRPOQOK 120
DB 77 RLQVPVEANVMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSLRPOQOK 136

QY 121 FPEALRGCEQEDSDIAFLIDGSGSIIPDFRMEKFEVSTVMEQKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCEQEDSDIAFLIDGSGSIIPDFRMEKFEVSTVMEQKSKTLFSLMQYSEEF 196

QY 181 RIHFTFKFQNNPNPRLVKPTIOLLGRTHRTATGVRKVIRELLINTNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNPRLVKPTIOLLGRTHRTATGVRKVIRELLINTNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVLPEDRSGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVLPEDRSGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQVN 316

QY 301 NFEALKTIONLREKIFAIEGTQTGSSSFHEHMQGSPSAITNSGPLLSTVSGYDNAG 360
DB 317 NFEALKTIONLREKIFAIEGTQTGSSSFHEHMQGSPSAITNSGPLLSTVSGYDNAG 376

QY 361 GVPLYSKESKPIFNTRVDSQNDAYLGYAAAILLRNVQSLVGLGAPYQHIGLVAMPR 420
DB 377 GVPLYSKESKPIFNTRVDSQNDAYLGYAAAILLRNVQSLVGLGAPYQHIGLVAMPR 436

QY 421 QNTGWMESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVLI GAPHYEQTRGGQSVCP 480
DB 437 QNTGWMESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVLI GAPHYEQTRGGQSVCP 496

QY 481 PRQORARWQDAVLYGEOQPMGRFGAALTVDLVNVDGKLTDAVAGPGEEDNRGAVILF 540
DB 497 PRQORARWQDAVLYGEOQPMGRFGAALTVDLVNVDGKLTDAVAGPGEEDNRGAVILF 556

QY 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLVRKAIMFNPREVARNVFCNDQVYKGEAGVRVCLHVOKSTRDLRREGQTSVVT 660
DB 617 PVLVRKAIMFNPREVARNVFCNDQVYKGEAGVRVCLHVOKSTRDLRREGQTSVVT 676

QY 661 YDLALDSGRPHSAVENETKNSRRTOVLGLOTCTETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSAVENETKNSRRTOVLGLOTCTETLKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFNLRLPVLAEADAQLFTALPFPEKNCNDNIQDDLSITFSMSLDCLVVG 780
DB 737 SLVGTPLSAFNLRLPVLAEADAQLFTALPFPEKNCNDNIQDDLSITFSMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGDSYRTQVTFEPFLDLSYRKVSTLQNRQSRWLACESASSTEV 840
DB 797 GPREFNVTVTRNDGDSYRTQVTFEPFLDLSYRKVSTLQNRQSRWLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKXANTYSENNPRTNKTFF 900
DB 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKXANTYSENNPRTNKTFF 916

QY 901 QLELPVKYAVNVTSHGVSTKYLNFPTASNTSRVMQHQYQVNSLQORSIPLSLVLFPV 960
DB 917 QLELPVKYAVNVTSHGVSTKYLNFPTASNTSRVMQHQYQVNSLQORSIPLSLVLFPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 PFGIQEERFATLKGNSLDFDWMYIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1080
DB 1037 PFGIQEERFATLKGNSLDFDWMYIKTSHNELLIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1096
QY 1081 KVPEFVEPNPLPIVGVSSVGLLALITAAALYKLGPFKQYKDMSEGPPGAEPO 1137
DB 1097 KVPEFVEPNPLPIVGVSSVGLLALITAAALYKLGPFKQYKDMSEGPPGAEPO 1153

RESULT 2

AA07360

ID AA07360 standard; protein; 1153 AA.

XX AC

XX AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha_d integrin;

XX KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

XX KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;

XX KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

XX KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

XX KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PP 16-NOV-1999; 99WO-US027139.

XX PR 16-NOV-1998; 98US-00193043.

XX PR 08-JUL-1999; 99US-00350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin NW, Van Der Vieren M;

XX DR WPI; 2000-387751/33.

XX PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
XX PT macrophage infiltration and reduce inflammation at central nervous system
XX PT injury sites.
XX PS Example 5; Fig 1; 270pp; English.
XX CC Integrins are a class of membrane-associated molecules that participate
XX CC in cellular adhesion. Integrins are made up of an alpha subunit and a
XX CC beta subunit. One class of human integrins are restricted to expression
XX CC in white blood cells and have a common beta2 subunit: the leukocyte
XX CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
XX CC have an important role in immune and inflammatory responses. The present
XX CC protein sequence is the human integrin alpha subunit CD11b. This sequence
XX CC was used in an alignment to identify a novel beta2 integrin alpha
XX CC subunit: alpha d (AAA60014 and AAB07359). The present sequence has
XX CC approximately 60% identity to the protein sequence of alpha d. The
XX CC Alpha d gene and protein may be useful in therapy for diseases linked to
XX CC atherosclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
XX CC (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the
XX CC inhibition of macrophage infiltration at the site of a central nervous
XX CC system injury. The monoclonal antibodies can also be used to detect and
XX CC diagnose Crohn's disease

QY	1021	FFGIOEFNATLKGNL	FDWYIKTSHNHLIVSTAEILFND	SVFTLLPGGCAFVRSQ	TET 1030
DB	1037	FFGIOEFNATLKGNL	FDWYIKTSHNHLIVSTAEILFND	SVFTLLPGGCAFVRSQ	TET 1036
QY	1081	KVEPFVFPNPLPLIV	GVSSVGGILLILALITAA	LKYLGFKKQYKDMSEGG	PPGAE PQ 1137
DB	1097	KVEPFVFPNPLPLIV	GVSSVGGILLILALITAA	LKYLGFKKQYKDMSEGG	PPGAE PQ 1153
RESULT 3					
AAU80252	AAU80252 standard; protein; 1153 AA.				
XX	AC	AAU80252;			
XX	DT	15-JUL-2002 (first entry)			
XX	DE	Human integrin 1 alpha-M subunit protein.			
XX	KW	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;			
XX	KW	inflammatory disease; autoimmune disorder; Crohn's disease;			
XX	KW	human immunodeficiency virus; HIV; myocardial infarction;			
XX	KW	Sjorgen's syndrome; rheumatoid arthritis.			
XX	OS	Homo sapiens.			
XX	XX	Key	Location/Qualifiers		
XX	XX	Misc-difference	499..500		
XX	XX	FT	/note= "Encoded by GGG CAG AGG"		
XX	XX	PN	WO200218583-A2.		
XX	XX	XX	07-MAR-2002.		
XX	XX	XX	31-AUG-2001; 2001WO-US027227.		
XX	XX	XX	01-SEP-2000; 2000US-0229700P.		
XX	XX	XX	(BLOO-) CENT BLOOD RES INC.		
XX	XX	XX	Springer TA, Shimoaka M, Lu C;		
XX	XX	XX	WPI; 2002-382964/41.		
XX	XX	XX	N-PSDB; ABK50045.		
XX	XX	XX	Modified integrin-I or integrin I-like domain polypeptide useful as an		
XX	XX	XX	immunogen to produce antibodies specific to polypeptide, comprises a		
XX	XX	XX	disulfide bond such that polypeptide is stabilized in a desired		
XX	XX	XX	conformation.		
XX	XX	XX	Disclosure; Page 109-112; 112pp; English.		
XX	XX	XX	This invention relates to a modified integrin-I or integrin I-like domain		
XX	XX	XX	polypeptide comprising at least one disulfide bond so that the domain is		
XX	XX	XX	stabilised in a desired conformation. The polypeptide of the invention		
XX	XX	XX	may have antiinflammatory or immunosuppressive activities. The		
XX	XX	XX	polypeptides of the invention have an open conformation and are useful as		
XX	XX	XX	immunogens to produce antibodies that selectively bind to integrin I-		
XX	XX	XX	domain; and for identifying a modulator of integrin activity, or of		
XX	XX	XX	interaction of an integrin and a cognate ligand. The polypeptide of the		
XX	XX	XX	invention, or antibodies (preferably anti-1FA-1 antibody) is useful for		
XX	XX	XX	treating or preventing an integrin mediated disorder which is an		
XX	XX	XX	inflammatory or autoimmune disorder in a subject and for inhibiting the		
XX	XX	XX	binding of an integrin to a cognate ligand such as Crohn's disease,		
XX	XX	XX	nephritis; human immunodeficiency virus (HIV), myocardial infarction,		
XX	XX	XX	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic		
XX	XX	XX	composition comprising the peptide of the invention is useful for		
XX	XX	XX	treating an integrin mediated disorder in a subject. The polypeptides		
XX	XX	XX	and/or active or antigenic fragments are useful as reagents for diagnosis		
XX	XX	XX	of integrin-mediated disorders. The present sequence represents the human		
XX	XX	XX	integrin-1 alpha-M protein subunit used to generate the mutant		
XX	XX	XX	polypeptides of the invention		

XX	Sequence 1153 AA;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
----	-------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db	977	RLNQTVINDPQVTFSENLSTCHTKERLPSPHSDFLAELRKAPVWNCISIAVCQRIQCDIP	103
Qy	1021	PFQIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTASILLFNDVSFTLLPGOGAFVRSQTET	1080
Db	1037	PFQIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTASILLFNDVSFTLLPGOGAFVRSQTET	1096
Qy	1081	KVPEPEVPNPLPIVGVSSVGGILLALLIATAALYKLGFFKRYQKDMSEGGPPGAEPO	1137
Db	1097	KVPEPEVPNPLPIVGVSSVGGILLALLIATAALYKLGFFKRYQKDMSEGGPPGAEPO	1153
RESULT 4			
ABG61469			
ID	ABG61469	standard; protein; 1153 AA.	
XX	XX	ABG61469;	
XX	XX	27-AUG-2002 (first entry)	
XX	XX	Human Beta2 integrin alphaCD11b subunit.	
XX	XX	Beta2 integrin; alphasubunit; CD11c subunit; CD11b subunit; LAD;	
KW	KW	leukocyte adhesion deficiency; inflammatory response; diabetes;	
KW	KW	multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;	
KW	KW	inflammatory bowel disease; Crohn's disease; ulcerative colitis;	
KW	KW	immune complex alveolitis; leukaemia; ICAM-1; anti-inflammatory;	
KW	KW	intracellular cell adhesion molecule; vascular cell adhesion molecule;	
KW	KW	locomotor recovery; locomotor damage; locomotor impairment;	
KW	KW	autonomic dysfunction; sensory dysfunction; spinal cord injury.	
XX	OS	Homo sapiens.	
XX	XX	WO200230980-A2.	
XX	XX	18-APR-2002.	
XX	XX	15-OCT-2001; 2001WO-US032059.	
XX	XX	13-OCT-2000; 2000US-00688307.	
XX	XX	(ICOS-) ICOS CORP.	
XX	XX	Gallatin WM, Van Der Vieren M;	
XX	XX	WPI; 2002-463260/49.	
XX	XX	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor	
XX	XX	recovery, inhibiting locomotor damage, limiting locomotor impairment, or	
XX	XX	limiting autonomic and sensory dysfunction following spinal cord injury.	
XX	XX	Example 5; Page 191-194; 270pp; English.	
XX	XX	The invention relates to promoting locomotor recovery, inhibiting	
XX	XX	locomotor damage, limiting locomotor impairment, or limiting autonomic	
XX	XX	and sensory dysfunction following spinal cord injury by administering an	
XX	XX	anti-alpha-d (Beta2 integrin alpha2 subunit) monoclonal antibody to a	
XX	XX	spinal cord injury victim. The method also involves the use of a ligand	
XX	XX	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,	
XX	XX	vascular cell adhesion molecule). The method is useful for promoting	
XX	XX	locomotor recovery, inhibiting locomotor damage, limiting locomotor	
XX	XX	impairment, or limiting autonomic and sensory dysfunction following	
XX	XX	spinal cord injury. In particular, the spinal cord injury comprises	
XX	XX	compression of the spinal cord. The antibodies are also useful for	
XX	XX	reducing inflammation at the site of a central nervous system injury. The	
XX	XX	specification also details the identification of Beta2 integrin alphaD	
XX	XX	cDNAs and proteins, for use in raising the antibodies. Beta2 integrins	
XX	XX	are implicated in diseases such as LAD (leukocyte adhesion deficiency,	
XX	XX	inflammatory response, diabetes, multiple sclerosis, arthritis, graft	
XX	XX	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative	
XX	XX	colitis, immune complex alveolitis and leukaemia. The present sequence is	
XX	XX	a Beta2 integrin alpha subunit sequence included for comparison with the	
XX	XX	Beta2 integrin alphasubunit sequences	

QY	1021	FFGIQEENFATLKGSLDFDWWIKTSHNLLIVSTABILFNDSVFTLLPGCGAFVRSGQTET	1098
DB	1037	FFGIQEENFATLKGSLDFDWWIKTSHNLLIVSTABILFNDSVFTLLPGCGAFVRSGQTET	1099
QY	1081	KVEPEPVPNPPLLVGSSVGGLHLLALITAAALVKLGFFKRYQKDMSEGPPGAEPO	1137
DB	1097	KVEPEPVPNPPLLVGSSVGGLHLLALITAAALVKLGFFKRYQKDMSEGPPGAEPO	1153
RESULT 5			
ID	AA014428	standard; protein; 1153 AA.	
XX	AAO14428;		
AC	AAO14428;		
XX	03-MAY-2002 (first entry)		
DE	Integrin Mac-1 alpha subunit.		
XX	Mac-1; integrin alpha subunit; variant integrin inserted domain protein; open conformation; integrin related inflammatory disorder;		
KW	integrin related immunological disorder; rheumatoid arthritis; ischaemia;		
KM	reperfusion; hypovolemic shock; infarction; cerebral shock;		
KW	viral infection; cancer; gene therapy; vaccine;		
KM	bioactive agent screening.		
OS	Unidentified.		
XX	WO200204521-A2.		
PN	17-JAN-2002.		
XX	09-JUL-2001; 2001WO-US021805.		
PD	07-JUL-2000; 2000US-0216600P.		
XX	(CALY) CALIFORNIA INST OF TECHNOLOGY.		
PR	(BLOO-) CENT BLOOD RES.		
PA	Springer T;		
XX	MPI; 2002-148167/19.		
DR	New integrin I domain protein having alteration in at least 2 noncontiguous regions and exits in an open conformation, useful for treating, preventing or suppressing inflammatory or immunological disorders.		
PT	Example 1; Fig 1F; 90pp; English.		
PS	The invention comprises structurally biased variant integrin inserted (I) domain proteins, wherein the alterations to the protein occur in at least two noncontinuous regions. Specifically the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating; ischaemia/reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin		
XX	Sequence 1153 AA;		
QY	Query Match	99.9%; Score 5868; DB 5; Length 1153;	
DB	Best Local Similarity	99.6%; Pred. No. 0;	
	Matches 1133; Conservative	3; Mismatches	1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLGGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSNTYVKGCLFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSNTYVKGCLFPGSNLRQOPQK 136
QY 121 FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTWMEOLKSKTLFSLMOYSEEF 180
Db 137 FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTWMEOLKSKTLFSLMOYSEEF 196
QY 181 RHFTPFKEQNNPNRSLVKPITQLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RHFTPFKEQNNPNRSLVKPITQLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 256
QY 241 TQGEKPGDPLGYEDVITPEADRGVRYVVLGVGDAFRSEKSRORBLTIASKPRDHVFOVN 300
Db 257 TQGEKPGDPLGYEDVITPEADRGVRYVVLGVGDAFRSEKSRORBLTIASKPRDHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVLPLYSKSKSTPINNTRVDSNDMDAYLGYAAAILLRNRVQSLVLCAPRYOHIGLVAMFR 420
Db 377 GVLPLYSKSKSTPINNTRVDSNDMDAYLGYAAAILLRNRVQSLVLCAPRYOHIGLVAMFR 436
QY 421 QNTGMVESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYETRGQGVSVCP 480
Db 437 QNTGMVESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYETRGQGVSVCP 496
QY 481 PRQQRARWQDAVLYGEOGPWGRFGAALTVLGDVNGDKLTVAICAPGEENRGAAYLIF 540
Db 497 PRQQRARWQDAVLYGEOGPWGRFGAALTVLGDVNGDKLTVAICAPGEENRGAAYLIF 556
QY 541 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGGQDLTWGGLVLTVCAGHVLRLRQ 600
Db 557 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGGQDLTWGGLVLTVCAGHVLRLRQ 616
QY 601 PVLVRKAIMENPREVARNVFECNDQVYKGEAGVRVCLHVOKSTRDLREGQIOVSVT 660
Db 617 PVLVRKAIMENPREVARNVFECNDQVYKGEAGVRVCLHVOKSTRDLREGQIOVSVT 676
QY 661 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRANP 720
Db 677 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRANP 736
QY 721 SLVGTPLSAFENLRPVLAEADQRLFTALPFERKNCNDNQCDDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFENLRPVLAEADQRLFTALPFERKNCNDNQCDDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLVSRKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPPLDLVSRKYSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMERTKNTBF 900
Db 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMERTKNTBF 916
QY 901 QLELPVKIAYVMVTSVSTKYNLFTASENTSRVMQHOVQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKIAYVMVTSVSTKYNLFTASENTSRVMQHOVQVSNLQORSPLISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVQCRIOCDIP 1020
Db 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVQCRIOCDIP 1036
QY 1021 PFGIOEBFNATLKGNLGDFWYIKTSHNELLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1080
Db 1037 PFGIOEBFNATLKGNLGDFWYIKTSHNELLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1096
QY 1081 KVEPFEVNPPLPLTVGSSVGGLLILLALITAAALYKLGFFKQYKDMWSEGGPPGAE 1137

Db 1097 KVEPFEVNPPLPLTVGSSVGGLLILLALITAAALYKLGFFKQYKDMWSEGGPPGAE 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
antiarthritic; immunosuppressive; antidiabetic; antichryoid;
neuroprotective; hinge region; immunoglobulin heavy chain;
CH2 constant region; CH3 constant region; IgG1;
antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
rheumatoid arthritis; myasthenia gravis; Grave's disease;
type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PF 25-JUL-2002; 2002US-00207655.
XX
PR 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENE-CRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
XX WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
subject having or suspected of having a malignant condition or a B-cell
disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclosure; SEQ ID NO 176; 157pp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
comprising a binding domain polypeptide that is fused to an
immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CH2 constant region polypeptide that is fused to the hinge region
polypeptide, and an immunoglobulin heavy chain CH3 constant region
polypeptide that is fused to the CH2 constant region polypeptide. The
hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
region polypeptide, derived from (a) having 3 or more cysteine residues;
where the mutated human IgG1 immunoglobulin hinge region polypeptide
contains 2 cysteine residues, where the first cysteine is not mutated; a
mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
(a) having 3 or more cysteine residues, where the mutated human IgG1
immunoglobulin hinge region polypeptide contains no more than one
cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
polypeptide, derived from (a) having 3 or more cysteine residues; where
the mutated human IgG1 immunoglobulin hinge region polypeptide contains
no cysteine residues. The binding domain-immunoglobulin fusion protein is
capable of at least one immunological activity comprising antibody
dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
binding domain polypeptide is capable of specifically binding to an
antigen. Also included are an isolated polynucleotide encoding the
binding domain-immunoglobulin fusion protein, a recombinant expression
construct comprising the polynucleotide (operably linked to a promoter),
a host cell transformed or transfected with a recombinant expression
construct, producing the binding domain-immunoglobulin fusion protein, a

FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1022..1024	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1045..1047	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1051..1053	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1076..1078	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1106..1134	
FT	Region	/label= putative transmembrane region	
PN	BP364690-A.		
XX	25-APR-1990.		
XX	17-AUG-1989;	89EP-00115159.	
XX	23-AUG-1988;	88US-00235353.	
PR	09-MAR-1989;	89US-00321239.	
XX	(DAND) DANA FARBER CANCER INST INC.		
PA	Springer TA, Corbi A;		
XX	WPI; 1990-125938/17.		
XX	N-PSDB; AAQ04043.		
DR	New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating		
DR	inflammation and viral infections, and in diagnosis.		
PT	Disclosure; Page ?; -pp; English.		
XX	Mac-1 alpha subunit is involved in the response to inflammation, i.e.		
XX	recognition of and migration to sites of inflammation. It also attaches		
CC	to cellular substrates as part of this function making it useful in		
CC	visualising endothelial tissue. Mac-1 is a member of the Integrin Gene		
CC	superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25		
CC	-MAR-2003 to correct PA field.)		
XX	Sequence 1153 AA;		
SQ	Query Match	99.78; Score 5858; DB 2; Length 1153;	
	Best Local Similarity	99.66; Pred. No. 0;	
	Matches 1132; Conservative	3; Mismatches 2; Indels 0; Gaps 0;	
Qy	1	FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAANQKSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAANQKSLYQCDYSTGSCPEI	76
Qy	61	RLOVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCEYVTKGLCFGLGSLNRQOPQK	120
Db	77	RLOVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCEYVTKGLCFGLGSLNRQOPQK	136
Qy	121	PPEALRGCPQSDSIAFLIDGSGIIPHDFFRMKEFVSTVMEQLKSKTILFSLMOYSEEF	180
Db	137	PPEALRGCPQSDSIAFLIDGSGIIPHDFFRMKEFVSTVMEQLKSKTILFSLMOYSEEF	196
Qy	181	RIHFTPEFQNNPNRSLVKPDIQTLGRTHATGVRVIRELLNITGARKNAFKILIVI	240
Db	197	RIHFTPEFQNNPNRSLVKPDIQTLGRTHATGVRVIRELLNITGARKNAFKILIVI	256
Qy	241	TDGKFGDPLGVEDVPEADREGVIRVIGVDGAFRSEKSRQELNTIASKPPRDHVPQVN	300
Db	257	TDGKFGDPLGVEDVPEADREGVIRVIGVDGAFRSEKSRQELNTIASKPPRDHVPQVN	316
Qy	301	NFEALKTIQNLREKI FAIEGTQGTGSSSPHEMSQGFSAATNSGPLLSTVGSYDWAG	360
Db	317	NFEALKTIQNLREKI FAIEGTQGTGSSSPHEMSQGFSAATNSGPLLSTVGSYDWAG	376
Qy	361	GVFLYTSKSTPTNMTVDSDMDAYLGVAALILRNVRQSLVGLGAPRYCHIGLVAMFR	420
Db	377	GVFLYTSKSTPTNMTVDSDMDAYLGVAALILRNVRQSLVGLGAPRYCHIGLVAMFR	436
Qy	421	QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVCP	480
Db	437	QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVCP	496
Qy	481	PRGQARWQCDVAVLYGEGQGPWGFAGALTIVLGDVNGDKLTDAVIGAPGEDNRCAYVLF	540
Db	497	PRGQARWQCDVAVLYGEGQGPWGFAGALTIVLGDVNGDKLTDAVIGAPGEDNRCAYVLF	556
Qy	541	HGTSGSISPSHESQRIAGSKLSPRLQVPGQSLSGQDLTMDGLVDLTVGAQGHVLLARSQ	600
Db	557	HGTSGSISPSHESQRIAGSKLSPRLQVPGQSLSGQDLTMDGLVDLTVGAQGHVLLARSQ	616
Qy	601	PVLRVKALMEFNPFEVARNVFECDNDQVVKGEARVRLVQKSTRDLREGQIQSVVT	660
Db	617	PVLRVKALMEFNPFEVARNVFECDNDQVVKGEARVRLVQKSTRDLREGQIQSVVT	676
Qy	661	YDLALDSGRPHSRVAFNETKSTRQTVLGTCTETKLQLPNCIEDPVSPIVLRNF	720
Db	677	YDLALDSGRPHSRVAFNETKSTRQTVLGTCTETKLQLPNCIEDPVSPIVLRNF	736
Qy	721	SLVGTFLSAFGNLRPVLAEDAQRLLFTALFPPEKKNCGNDNICODDLSITFSPMSLCLVVG	780
Db	737	SLVGTFLSAFGNLRPVLAEDAQRLLFTALFPPEKKNCGNDNICODDLSITFSPMSLCLVVG	796
Qy	781	GPREFNVTVVNDGEDSYRTQVTFPDLISYKRVSTLONQORSORSNRLACESASSTEV	840
Db	797	GPREFNVTVVNDGEDSYRTQVTFPDLISYKRVSTLONQORSORSNRLACESASSTEV	856
Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSNNMPTNKTEF	900
Db	857	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSNNMPTNKTEF	916
Qy	901	QLELPVKYAVVTVSHGVSTKYLNFETASENTRVMQHQYOVNSLQORSLPISLFLVLPV	960
Db	917	QLELPVKYAVVTVSHGVSTKYLNFETASENTRVMQHQYOVNSLQORSLPISLFLVLPV	976
Qy	961	RLNQTVINDRPOVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCSTAVQRIQCDIP	1020
Db	977	RLNQTVINDRPOVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCSTAVQRIQCDIP	1036
Qy	1021	FFGIOEEFNATLKMSLSPDWYIKTSHNHLIVSTAEILFENDSVFTLLPCOGAPVRSQTET	1080
Db	1037	FFGIOEEFNATLKMSLSPDWYIKTSHNHLIVSTAEILFENDSVFTLLPCOGAPVRSQTET	1096
Qy	1081	KVEPFEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEQ	1137
Db	1097	KVEPFEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEQ	1153
RESULT 8			
AA07120			
ID	AA07120	standard; protein; 1163 AA.	
XX	AA07120;		
AC	25-MAR-2003	(revised)	
DT	05-FEB-1991	(first entry)	
XX	p150.95	alpha subunit encoded by clone lambdaX47.	
XX	p150.95	leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;	
XX	rhinovirus.		
XX	Synthetic.		
XX	Key	Location/Qualifiers	
FT	Region	1..19	
FT	Region	/label= signal peptide	
FT	Region	20..44	
FT	Region	/label= N-terminus	
FT	Modified-site	61..63	
FT	Modified-site	/label= glycosylation site	

FT	Modified-site	89..91	/label= glycosylation site	Db	198	QTHFTFEFRRTSNPLSLASVHQLQFTYTATAIQNVVHRLPHASGARDATKILIVI	257
FT	Modified-site	385..387	/label= glycosylation site	Qy	241	TDGEKCDPLGVEDVIPEADREGVIYVIGVDAFSEKSEKQELNITIASPPRDHVFQVN	300
FT	Modified-site	392..394	/label= glycosylation site	Db	258	TDGKKGDSLDYKDVIPMDAAGLIIYAIGVLAQFNRSWKELNDIASPSQEHIFKVE	317
FT	Modified-site	697..699	/label= glycosylation site	Qy	301	NFRALKTIQNLREKIPAEIGTGTGSSSSPEHEMSQEGFSAAITNSGSLTSTVGSYDWAG	360
FT	Modified-site	735..737	/label= glycosylation site	Db	318	DFDALKDIOQLKPKIPAEIGTGTGSSSSPELEMAQEGFSAVTPDGPVLGAVGSFTWSG	377
FT	Modified-site	899..901	/label= glycosylation site	Qy	361	GVFLYTSKESKSTINMTFRVDSMDNDAYGLYAAAILLNRVQSLVGLGAPRYCHILGAMER	420
FT	Modified-site	904..906	/label= glycosylation site	Db	378	GAFLYPNMSPFTINMSQENVMDRDSYLGYSTELALWKGVQSLVGLGAPRYCHILGAMER	437
FT	Modified-site	939..941	/label= glycosylation site	Qy	421	QNTGMESNANVAGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP	480
FT	Modified-site	1050..1052	/label= glycosylation site	Db	438	QVSRQWRKAEVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEYQTRGGQVSVCP	497
FT	Modified-site	1108..1133	/label= glycosylation site	Qy	481	PRGORAWQCDVLYGEGQOPWGRFGAALTVLGDVNGCKLTVDVAIGAPGEDNREGAYLFP	540
FT	Domain		/label= transmembrane	Db	498	PRGWR-RWVCDVLYGEGQOPWGRFGAALTVLGDVNGCKLTVDVVIAGPGEENRGAYLFP	556
XX	WO9010646-A.			Qy	541	HGTSGSGISPSHSORIASGKSLSPLOYFGOSLGGQDLTMDGLVDLTVGAGQHYLLRSQ	600
XX	20-SEP-1990.			Db	557	HGVLGPSISPSHSORIASGQSSRELQYFGQALSGQDLTQDGLVDLAVGARGQVLLLR	616
PF	09-MAR-1990;	90WO-US001257.		Qy	601	PVLRTALIMBFNPREVARNVFECDQVYVKGKAGEVRVCLHVQKSTDRDLREGQIQSVVT	660
PR	09-MAR-1990;	90WO-US001257.		Db	617	PVLWVGVMQFIPAEIPRSAPFECEQVWVSEGTLLVQSNCLYIDKRSKSLGSRDLQSVT	676
XX	(DAND) DANA FARMER CANCER INST INC.			Qy	661	YDLALDSGRHSRAVFNETKSTERTQVGLGTQTCETLKLQLPNCIEDPVSPTVLRNLF	720
XX	Corbi AA, Springer TA;			Db	677	LDLALDPERLSPRATFQETKRSLSRVRLGLKACENFNLLFSCVEDSVTPITLRNLF	736
DR	WPI; 1990-304985/40.			Qy	721	SLVGTPLSAFGLNPLVLAEDAQRFTALFPPEKNCNDNICQDDLSTTFSPMSDCLVVG	780
DR	N-PSDB; AAQ06068.			Db	737	TLVGKPLLAFLNLRPMLAALQRYETASLPFKNCGADHICQDNLGISFSFPGLKSLVG	796
PT	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-			Qy	781	GPREFVTVVRNDGDSYRTQVTFPPPLDLSYKVKSTLQNRQSORSWRLACESASSTEV	840
PT	unit of p150.95 cell surface adhesion receptor, opt. together with a beta			Db	797	SNLELNAEVWVWMDGDSYGTITITFHPAGLSYRYVAGGQKQQLRSUHTCDSA--PVA	854
XX	chain of CD-18 family.			Qy	841	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	900
XX	Disclosure; Fig 3; 59pp; English.			Db	855	SGQWTSISCRINHILFRGGAQITELATFDVSPKAVLGDRLLLTANVSSENTPRTSKTTF	914
CC	Clone lambda X47 was isolated from a cDNA library constructed from total			Qy	901	QLELPVKYATVMTVSHGVSTKYLNTAS-RNTRVWQHGYOVNLSGORSPLSLVPLVP	959
CC	RNA extracted from phorbol myristate acetate stimulated HL-60			Db	915	QLELPVKYATVMTVSHGVSTKYLNTAS-RNTRVWQHGYOVNLSGORSPLSLVPLVP	974
CC	myelomonocytic cells. The library was screened with oligonucleotide			Qy	960	VLNQTVIWDPRQVTFSENLSSTCTKTERLPSSHDFLAELKAPVWVNCISVACORIQCDI	1019
CC	probes based on tryptic peptide fragments of p150.95. The sequence can be			Db	975	VELNQEAWMVDVEVSHPCNPSPSRCSQKIAPPASDPLAHIQKNPVLDCSIAAGCLRFCDV	1034
CC	attached to appropriate control elements and expressed in prokaryotic and			Qy	1020	PPFGIQEEFNATLKNLSFDWYIKTSHNHLILIVTABIILFNDSTVFTLLPGQAPVRSQTE	1079
CC	eukaryotic cells. The protein can be used to treat or prevent rhinoviral			Db	1035	PSFSVQEBLDTLAKNLISFGWVRQILQKKSVSWVAEITFTDTSVYSQLPQGERFMAQTT	1094
CC	infection because it interacts with ICAM-1 and inhibits cell-virus			Qy	1080	TKVPEPPEVNPDLPLTVGSSVGLLALLALITAAALYKLGPFKQYKDMSE	1128
CC	attachment. It can also be used as an anti-inflammatory agent. See also			Db	1095	TVLEKRYKVNHPPLIVGSSIGALLLALITAVLTKVGFKKQYKEMEE	1143
CC	AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003						
CC	to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)						
XX	Sequence 1163 AA;						
SQ	Query Match	59.0%; Score 3464; DB 2; Length 1163;					
	Best Local Similarity	61.1%; Pred. No. 2.5e-279;					
	Matches 690; Conservative 138; Mismatches 295; Indels 6; Gaps 4;						
Qy	1 FNLDENAMFFQBNARGFGQSVVQLQGRVWVGAPOEIVAAVQGRSLYQDYSTGSCBPI	60					
Db	20 FNLDTELTAFRVDAGFGDSVQVYANVWVGAPOKITAANQTCGLYQCYSTGACBPI	79					
Qy	61 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQPOK	120					
Db	80 GLOVPPEAVNMSLGLSLAATTSPPOLLACGPTVHQCSENTYVKGCLFLFGSNLRQPOK	137					
Qy	121 PFEALRGCPQEDSDIAPLDGSGSIIPHDFRNKEFVSTVMEQLKSKTLFSLMYSEFP	180					
Db	138 LPVSRQECPRQEDIVFLIDGSGSISSRNFAVNFVRAVISQFQRPSTQPSLMQFSNKF	197					
Qy	181 RIHFTKGFQNNPNSLKPITOLLGRTHATGVKVRBELNITNGARKNAFKILIVI	240					

RESULT 9
AAW65091
ID AAW65091 standard; protein; 1163 AA.
XX
AC AAW65091;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11c subunit protein.
XX

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 XX US728533-A.
 XX 17-MAR-1998.
 XX 07-JUN-1995; 95US-00485618.
 XX 23-DEC-1993; 93US-00173497.
 XX 05-AUG-1994; 94US-00286889.
 XX 21-DEC-1994; 94US-00362652.
 XX (ICOS-) ICOS CORP.
 XX Van Der Vieren M, Gallatin WM;
 XX WPI; 1998-206565/18.
 XX Screening assay for modulators of integrin binding - using immobilised or
 XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.
 XX Example 5; Fig 1A-D; 106pp; English.
 XX This sequence represents a human beta-integrin CD11c subunit which is
 XX used to describe a method for identifying compounds that modulate the
 XX interaction of the beta-integrin alpha-d subunit with a binding partner
 XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
 XX -d binding partner, one of which is immobilised and the other of which is
 XX labelled, in the presence of a test compound, and determining if the
 XX compound affects binding between the alpha-d polypeptide and alpha-d
 XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 XX comprising the cytoplasmic, transmembrane or extracellular domain of
 XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
 XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,
 XX asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 XX and rheumatoid arthritis
 XX Sequence 1163 AA:
 SQ Query Match 58.7%; Score 3450; DB 2; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 3.8e-278;
 Matches 689; Conservative 144; Mismatches 290; Indels -6; Gaps 4;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGSRVVGAPQEIIVAAHQSLGYQDYSTGSCPEI 60
 DB 20 FNLDTEELAFRVDSAGFSDSVVQYANSVWVGAPQKILAAQIGLYQGYSTGACPEI 79
 QY 61 RLQVPEAVNMSGLSLAATTSPPQLACGPTVHQCSENYTKYKGLCFPLGSLRQPOPK 120
 DB 80 GLQVPEAVNMSGLSLAATTSPPQLACGPTVHQCSENYTKYKGLCFPLGSLRQPOPK 137
 QY 121 FPEALRGCEDESDIAFLDGSGSIIPHPFRMKFVPTVMEQLKSKTFLFSLMYSSEF 180
 DB 138 LPVSRCECPQEQDIFVLIDGSGSIISRRNPATMNPRAVISQFQFSTQSLMQSNKF 197
 QY 181 RIHFTKBFQNNPNSLVKPIETQLGRTHATGVRKVRRELLININGARKNAFKILIVI 240
 DB 198 QTHFTFEFPRTNPLSLASVHLQCFVTATAIQNVVHRLFHASYGARRDAIKILIVI 257
 QY 241 TDCEKGDPLGYEDVPEADREGVIRVIGVGAFRSEKSRQBLNTIASKPRDHVPQVN 300
 DB 258 TDCKKGDSDYNDVPMADAAGIIRYVIGVGLAFQNRNSWKELNDIASKPSHELFKYE 317
 QY 301 NFELAKTIQNLREKIFAIBGTQTGSSSPSEHEMSQEGFSAATISNGPLSTVGVSDWAG 360
 DB 318 DFDALDIQNLREKIFAIBGTQTGSSSPSEHEMSQEGFSAATISNGPLSTVGVSDWAG 377
 QY 361 GVFLYTSKESKSTINMTRVDSMDNDAVLGYAAAILNRRVQSLVGLAPRVQHIGLVAMFR 420

DB 378 GARLYPENMSPFFINMSQENVDMRDSYLGYSITELALWKGVQLGAPRYOHIKAVIFI 437
 QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQGVSVCP 480
 DB 438 QVSRQWRKAEVIGTIGSTFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQGVSVCP 497
 QY 481 PRGORARQCDVAVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGBEDNKGAVYLF 540
 DB 498 PRGWR-RWMCDAVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGBEDNKGAVYLF 556
 QY 541 HGTSGSGISPSHSORLAGSKLSPLOYFGQSLGSGQDLTMDGLVLTVAAGHVLILRSQ 600
 DB 557 HGVLGPSISPSHSORLAGSKLSPLOYFGQSLGSGQDLTMDGLVLTVAAGHVLILRSQ 616
 QY 601 FVLRVKAIMBEFNPREVARNVPCNDQVYVKGKAGEVAVCLHVQKSTRDLREGIQSVVT 660
 DB 617 FVLVGVSMQFIPAEIIPSAFECREQVVSQETLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRTOVQLGLTQTCETKLQKLPNCLEDPVSPVLRNF 720
 DB 677 LDALAPGRUSPRAIFOETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRN 736
 QY 721 SLVGTSLSAFQNLPRVLAEDAQRLLFTALFPPEKNCNDNICODDLISITFSFMSLDCLVVG 780
 DB 737 TLVGKPLAFNLPRVLAEDAQRLLFTALFPPEKNCNDNICODDLISITFSFMSLDCLVVG 796
 QY 781 GPRFRNVTVVRNDGEDSYRTQVTPFPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
 DB 797 SNLEINAEVWVWVNDGEDSYRTQVTPFPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 854
 QY 841 SGALKSTSCSNHPIFPENSVETNIPTFVDVSKASLGNKLLKXNTSENNMPTNKTEF 900
 DB 855 SQGTWSTSCRINHLIFRGGAQITFLATFVSPKAVGLDRLLLIANVSENIPRTSKIF 914
 QY 901 QLELPKVAIVYVTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVNLGORSPLISLVFLVP 959
 DB 915 QLELPKVAIVYVTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVNLGORSPLISLVFLVP 974
 QY 960 VELNCTVMDRPOVTFPSNLSSTCTKRLPSHSDFLAELRKAPVNCISVACORIQDI 1019
 DB 975 VELNCTVMDRPOVTFPSNLSSTCTKRLPSHSDFLAELRKAPVNCISVACORIQDI 1034
 QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPQOGAFVRSQTE 1079
 DB 1035 PFSVQBELDFTLKNLSFDWYIKTSHNHLIVSTAEILFNDSTVFTLLPQOGAFVRSQTE 1094
 QY 1080 TKVEPFPVNPPLPLIVGSSVGLILLALITATYLYKLGFFKEQYKDMSE 1128
 DB 1095 TVLEKYKVNPIPLIVGSSVGLILLALITATYLYKLGFFKEQYKDMSE 1143
 RESULT 10
 AAB07361
 ID AAB07361 standard; protein; 1163 AA.
 XX AAB07361;
 XX 17-JAN-2001 (first entry)
 XX Human CD11c protein sequence.
 DE Human; macrophage infiltration inhibition; alpha_d integrin;
 KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; IAD; Type 1 diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11c.
 XX Homo sapiens.
 OS WO200029446-A1.
 XX PN
 XX

PD 25-MAY-2000.
 XX
 PF 16-NOV-1999; 99WO-US027139.
 XX
 PR 16-NOV-1998; 98US-00193043.
 PR 08-JUL-1999; 99US-00350259.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gallatin MW, Van Der Vieren M;
 XX
 DR WPI, 2000-387751/33.
 XX
 XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous system
 PT injury sites.
 XX
 XX Example 5; Fig 1; 270pp; English.
 XX
 CC Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This sequence
 CC was used in an alignment to identify a novel beta2 integrin alpha
 CC subunit: alpha d (AA60014 and AA607359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha.d. The
 CC alpha.d gene and protein may be useful in therapy for diseases linked to
 CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
 CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
 CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
 CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
 CC inhibition of macrophage infiltration at the site of a central nervous
 CC system injury. The monoclonal antibodies can also be used to detect and
 CC diagnose Crohn's disease
 XX
 SQ Sequence 1163 AA;
 Query Match 58.7%; Score 3450; DB 3; Length 1163;
 Best Local Similarity 61.0%; Fred. No. 3.8e-278;
 Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;
 QY 1 FNLDPTENAMTFOENARGGQSVVQLQSGSRVVGVAQPEIVAAHQESLYCCDYSGSCBPI 60
 DB 20 FNLDTEELTAPEVDSAGDSVVQYANSMWVGAPQKLIIRANQIGLYCCGYSTGACBPI 79
 QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHTCTSENVTYKGLCLPLFSNLRQOPQK 120
 DB 80 GLOVPEAVNMSLGLSLASTTSPOLLACGPTVHTCTSENVTYKGLCLPLFSNLRQOPQK 137
 QY 121 PPEALRGCPQEDSDAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLPSLMOYSBEF 180
 DB 138 LPSVQCEPRQEDIVFLIDGSGSISSRNFAFMNFAVISOQRPSTQSLMOFSNKP 197
 QY 181 RHFTFKFQNNPNRPSLVKPTQLLGRTHTAAGVKRVIRELLNITNGARKNAFKILIVI 240
 DB 198 QTHFTFEFRNTSNPLSLASVHQLQGGFTYATATQNVVHRLFAASYGARRDAIKILIVI 257
 QY 241 TDGKPGPLGVEDVTPADREGVIRYVIGVGDAPFSEKSKQELNTIAKPPRDRHPOVN 300
 DB 258 TDGKKEGSLDYKQVIPHADAGIIRYAVGLAFQNRNSWKELNDIAKSKSQEHFKVE 317
 QY 301 NFEALKTIQNRKIFAEIGTQSSSFEHMSQEGFSAATSGNPLLTGVSQYDWAQ 360
 DB 318 DFDALDKIQNLKEKIFAEIGFETISSSSFELEVAQEGFSAVFTPDGPVLGAVGSFTWSG 377
 QY 361 GVPLTSEKSTFTNTRVDSMDNAYLYGAAATILRNVSQSLVLCAPRYOHIGLVAMFR 420
 DB 378 GAFLYPPNMSPTFINNSQENVDMDRSYLGSTELALWKGVSQSLVLCAPRYOHIGKAVIFI 437
 QY 421 QNTGMWESNANVKGQIQGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQRGGQSVCP 480

Db 438 QVSRQRMKAEVIGTQIGSYFGASLCSVDVDTDGSSTDLVLIGAPHYYEQRGGQSVCP 497
 QY 481 PRGORARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAYILF 540
 Db 498 PRGWR-RWNCDAVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVVGAPGEBENRGAYILF 556
 QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGOSLGGDLTMDGLVTLVGAQGHVLLRSQ 600
 Db 557 HGVLPFSISPSHSORTAGSLSRLQYFGQALSGGDLTQDGLVTLVGAQGHVLLRTR 616
 QY 601 PVLRFVKAIMEFNEPREVARNYFECNDQVVKGEAGEVRVCLHVQKSTFDRLEBEOIQSVVT 660
 Db 617 PVLVGVSMQFIPAEIPRSFAFECREQVWSQTLVQSNICLYIDKRSKLLSGSRDLQSSVT 676
 QY 661 YDLALDSGRPHSAVFNENKSTRTQVIGLQTCTETLKLQPCNCTEDVSPVTVIRNF 720
 Db 677 LDALAPGRSLPRAIQETKRSLSRVRLGKAHCENFNLLPSCVEDSVIPIILRNF 736
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALPFPEKNCNGNDNICQDDLSITPFMSLDCIIVG 780
 Db 737 TLVGKPLAFLNLRPMLAALAQRYFTASLPFEKNCQADHICQDNLGISFSPGLKSLVG 796
 QY 781 GPREFNVTVRNDGEDSYTQVTPFPDLSTRKYSTLQNRQSORSWRLACESASTEV 840
 Db 797 SNLELNAEVMVMNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
 QY 841 SGALKTSCSINPIRPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTPEF 900
 Db 855 SQTWSTSCINHLIFRGQAQIFLAFDVPSPKAVGLDRLLANVTSENMPRTNKTPEF 914
 QY 901 QLELPVYAVVYVWVTSKVSQKYNLTAS-ENTSRVNOHQYQVSNLQGRSLPISLVLPV 959
 Db 915 QLELPVYAVVYVWVTSKVSQKYNLTAS-ENTSRVNOHQYQVSNLQGRSLPISLVLPV 974
 QY 960 VRLNQTIVDRPQVTFSENLSSCTCHTKERLPSSDFLAELRKAPVNCSTAVCORICDI 1019
 Db 975 VELNQEAVMVMDEVSHPFNPSLCSSEKIAPPASDFLAHQKQVLPDSCSAGCGRFCDV 1034
 QY 1020 PFGIQEENATLKENLSDFWYIKNTSHNLLIYVSTABIFENDSVFTLLPGCAFVRSQTE 1079
 Db 1035 PSFSVQSELDFTLKENLSPGVRQILQKKYVSVVABIIEDTSVYSQLPQGEAFMRAQTI 1094
 QY 1080 TKVPPPEVNPPLTVGSSVSGELLALITAAALYKLGFFKRYKDMXSE 1128
 Db 1095 TVLEKYKHNPPIPLTVGSSIGGLLALITAVLYKVGFFKRYKEMMBE 1143
 RESULT 11
 ABG61470
 ID ABG61470 standard; protein; 1163 AA.
 XX
 AC ABG61470;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human Beta2 integrin alphaCD11c subunit.
 XX
 KW Beta2 integrin; alpha d subunit; CD11c subunit; CD11b subunit; LAD;
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX
 FN WO2002030980-A2.
 XX
 PD 18-APR-2002.
 XX

PF	15-OCT-2001; 2001WO-US032059.	DB	438	QVSRQWRKAEVIGTQIGSYFGASICSVDVDTGSDTDLVLIGAPHYRQTRGGQSVSCPL	497
XX					
PR	13-OCT-2000; 2000US-00688307.	QY	481	PRGQARWQCDVLYGEGQWPGFAGALTVDGWDGKLTVDVAGDEENRGAUYLF	540
XX					
PA	(ICOS-) ICOS CORP.	DB	498	PRGMR-RWCDVLYGEGQWPGFAGALTVDGWDGKLTVDVAGDEENRGAUYLF	556
XX					
PI	Gallatin WM, Van Der Vieren M;	QY	541	PGTSGGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQHVLRLRSQ	600
XX					
DR	WPI; 2002-463260/49.	DB	557	HGVLGPSISPSHSQRIAGSLSRLQYFGQALSGQDLTQDGLVDLAVGARGOVLLATR	616
XX					
XX	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor	QY	601	PVLRVKAIMEPNRPREVARNFECNDQVYKGKAGRVRVCLHVQKSTRDELREGQIOSVUT	660
PT	recovery, inhibiting locomotor damage, limiting locomotor impairment, or				
PT	limiting autonomic and sensory dysfunction following spinal cord injury.	DB	617	PVLWVGVMQFIPABIPRSAFECEQVYSEQTLVQSNICLYDKRKSLLSRDQSSVT	676
XX					
PS	Example 5; Page 194-198; 270pp; English.	QY	661	VDLALDSGRPHSRVAFNETKSTRTOVLTGLTCTETKLQLPNCIEDPVPVILRLNF	720
XX					
CC	The invention relates to promoting locomotor recovery, inhibiting	DB	677	LDLALAPGRLSPRALFOETKNSLSRVVLGAKHCENFLLPSCVEDSVIPIILRLNF	736
CC	locomotor damage, limiting locomotor impairment, or limiting autonomic				
CC	and sensory dysfunction following spinal cord injury by administering an	QY	721	SLVGTPLSAGNLRPVILABDAORLTALPPFKKNGNDNICODDLISIFSPMSLDCLVVG	780
CC	anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a				
CC	spinal cord injury victim. The method also involves the use of a ligand	DB	737	TLVHGKPLLAFLRMLPMLAALAQRYFTASLPFKKNGADHICODNLGIGSFPGLKSLLVG	796
CC	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,				
CC	vascular cell adhesion molecule). The method is useful for promoting	QY	781	GPREENVTVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV	840
CC	locomotor recovery, inhibiting locomotor damage, limiting locomotor				
CC	impairment, or limiting autonomic and sensory dysfunction following	DB	797	SMLENAEVMWWDGEDSYGTITFESHAGLSYRVABGQKQGLRSLHLTC--CSAPVG	854
CC	spinal cord injury. In particular, the spinal cord injury comprises				
CC	compression of the spinal cord. The antibodies are also useful for	QY	841	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVSENMPRTNKTBF	900
CC	reducing inflammation at the site of a central nervous system injury. The				
CC	specification also details the identification of Beta2 integrin alpha	DB	855	SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNTPTSKTIF	914
CC	cDNAs and proteins, for use in raising the antibodies. Beta2 integrins				
CC	are implicated in diseases such as LAD (leukocyte adhesion deficiency,	QY	901	QLELPVKYAVYVAVTSHGVSTKYLNTAS-ENTSRVMOHQYOVSNLQORSLPISLVFLVP	959
CC	inflammatory response, diabetes, multiple sclerosis, arthritis, graft				
CC	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative	DB	915	QLELPVKYAVYIVVSSHQFTKYLNFSSSEKESHVAVHRYQVNNLQORDLPVSNFWTFP	974
CC	colitis, immune complex alveolitis and leukaemia. The present sequence is				
CC	a Beta2 integrin alpha subunit sequence included for comparison with the	QY	960	VELNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI	1019
CC	Beta2 integrin alpha protein sequences				
XX		DB	975	VELNQEAVWMDVEVSHPPQPSLRCSSEKLIAPPASDPLAHIQKNPVLDCSIAGCLPRCDV	1034
SQ	Sequence 1163 AA;				
	Query Match 58.7%; Score 3450; DB 5; Length 1163;	QY	1020	PFPGTQEBFNATLKENLSFDWYIKTSHNHLILVSTAEILFNDVSFTLLRQGFVRSOTE	1079
	Best Local Similarity 61.0%; Pred. No. 3.8e-278;				
	Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;	DB	1035	PSFSVQEELEDFTLKGNLSFGWVQILQKVVSVVVAEILFDTSVYSQLPQGFAPMRAQTI	1094
QY	1 FNLDTENAMFQENARFGQSVVOLQGSRVVVGAPQIIVANQEGSLYQDYTGSCPEI	QY	1080	TKVVEPEVNPPLIVGSSVGGILLALITAAALYKLGFFKRYKCMME 1128	
DB	20 FNLDTBELTAFRVDGAGFSDVQYANVWVVGAPQKIIAANOIGGLYQCGYSTGACEPI	DB	1095	TVLEKYKVENPIPLIVGSSIGGLILLALITAVLYKVPFPKRYKEMME 1143	
QY	61 RLQVPVZAVNMSLGLSLAATTPPQLACGPTVHTQCTSENTYVKGLCFPLFGSNLRQOPQ				
DB	80 GLQVPEAVNMSLGLSLASTTSPQLACGPTVHCEGRNMYLTGLCFLLGPT--QLTOR				
QY	121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEPSTVMEQLKSKTFLPSLMQYSEEP				
DB	138 LFSVRQECPEQDQIVFLIDGSGSISRNFTAMNFVRAVISQFQRPSTQFSLMQFSNKF				
QY	181 RIHTFEKFNPNPRLSVKPIITOLLGRHTATGVRKVIRELNITNGARKNAFKILIVI				
DB	198 QHTTFEERTSNPLSLASVHQSQFTTATAIQNVHVELPHASYGARDALKILIVI				
QY	241 TDGKPGDPLGYEDVIPADREGVIRVYVGDAFRSEKRSQELNTIASKPRPDHVFQVN				
DB	258 TDGKKKGSDLDYKDVIPMADAAGIIRYVGLAFQNRNSWKELNDIASKPSQEHFKVE				
QY	301 NFEALKTQNLREKIPAIETQGTGSSSEHEMSQSGFSAITNSGPLLSTVGSYDWAG				
DB	318 DFDALDKDQLQNLKREKIPAIETETISSSFELEMAQEGFSAVTFPDGVLGAVGSFTWSG				
QY	361 GVFLYTSKSKSTFTINMTVDSDMDNDVIGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR				
DB	378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVSILVGLAPRYOHIGKAVIFI				
QY	421 QNTGWESNANVKTQICAYFGALCSVDVDSNSTDLVLIGAPHYRQTRGGQSVSCPL				

RESULT 12
ABU07406
ID ABU07406 standard; protein; 1163 AA.
XX ABU07406;
XX AC ABU07406;
XX DT 28-JAN-2003 (first entry)
XX DE Protein differentially regulated in prostate cancer #9.
XX KW Prostate cancer; gene expression; differential regulation;
XX KW molecular marker; drug target; cancer detection; cancer diagnosis;
XX KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX OS Homo sapiens.
XX PN WO200281638-A2.
XX PD 17-OCT-2002.
XX PF 08-APR-2002; 2002WO-US010824.
XX PR 06-APR-2001; 2001US-0281731P.
XX PR 06-APR-2001; 2001US-0281732P.
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Sun Z, Jay G;
XX DR WPI; 2003-058520/05.
XX PT Novel genes which are differentially regulated in prostate cancer, useful
XX PT for diagnosing prostate cancer in prostate tissue sample and assessing
XX PT therapeutic or preventive intervention in prostate cancer patients.
XX PS Claim 1; Page 225-228; 416pp; English.
XX CC The invention describes genes (I) which are differentially regulated in
XX CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX CC sample comprising prostate tissue, which involves determining the number
XX CC of target genes which are differentially-regulated in the sample, where
XX CC the number is indicative of the probability that the sample comprises
XX CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX CC intervention in a subject having a prostate cancer, which involves
XX CC determining the expression levels in a sample comprising prostate tissue
XX CC of target genes which are differentially-regulated in prostate cancer.
XX CC Preferably, the expression levels of at least 10 genes are determined.
XX CC (I) is also useful for identifying agents that modulate a biological
XX CC activity of a polypeptide differentially-regulated in prostate cancer
XX CC cells, which involves contacting a polypeptide differentially-regulated
XX CC in prostate cancer cells with a test agent under conditions effective for
XX CC the test agent to modulate a biological activity of the polypeptide, and
XX CC determining whether the test agent modulates the biological activity. (I)
XX CC is useful as molecular markers, as drug targets, and for detecting,
XX CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
XX CC preventing or treating, determining predisposition to diseases and
XX CC conditions especially relating to prostate cancer. (I) and its expression
XX CC products are used in the diagnostic test to assay for presence of cancer
XX CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX CC of cancer, its stage of development, the nature of genetic defect, etc.
XX CC The polypeptide encoded by (I) can be used as target for therapy or drug
XX CC discovery. (I) can also be used for expressing the polypeptide and thus
XX CC in searching specific binding partners of the polypeptide. (I) is useful
XX CC in therapeutic applications to treat prostate cancer. The identification
XX CC of specific genes, and groups of genes, expressed in pathways
XX CC physiologically relevant to prostate cancer permits the definition of
XX CC functional and disease pathways and the delineation of targets in these
XX CC pathways which are useful in diagnostic, therapeutic, and clinical
XX CC applications. This is the amino acid sequence of a protein differentially
XX CC regulated in prostate cancer
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3448; DB 6; Length 1163;
Best Local Similarity 60.9%; Pred. No. 5.5e-278;
Matches 688; Conservative 136; Mismatches 299; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFOENARFGQSQVQLQGGRRVVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFDSVVQVANSVVVVGAPQKITAANTQGLYQCGYSTGACEPI 79
QY 61 RLQVPVEANMGLSLAATTPSPOLLACGPTVHOCSTENTYVKGCLFLGSLNRPQOK 120
DB 80 GLQVPEAVNMGLSLASTTSPOLLACGPTVHOCSTENTYVKGCLFLGSLNRPQOK 137
QY 121 FPEALRGCPQSDIDAFIDGSGSIIPDFRMRKPFVSTVMEQLKXSKTLFSLMOYSBEF 180
DB 138 LPVSRQECRQEQDIVFLIDGSGSISSRNFAFMNFRVAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHFTKPEFQNNPNRSLVKPTTOLLGRTHWTATGVRKVIRELINTNGARKNAFKILIVI 240
DB 198 QTHLTPEEFRTSNPLSLASVHQLQGTYYTATAIGNVVRHFLPHASYGARRDATTILIVI 257
QY 241 TDGEKFGDPLGVEDVTPRADRECVIRYVIGVDGAPRSEKSRQBLNTIASKPRDHVQVFN 300
DB 258 TDGKKGEDTLDYKDVIPMDADAGIRYALGVGLAFQNRNSWELNDIASKPSQEHFKVKE 317
QY 301 NFEALKTIONQUREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGVDMAG 360

DB 318 DFDALDKIQTOLREKIPFIECTETTTSSSFELEMAQEGFSAVFTPDGPVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLENRVQSLVLAGPRYQHTGLVAMPR 420
DB 378 GAFLYPPNNSPTFINMSQENVDNRDSYLGYSTELALWKGVSILVLAGPRYQHTGKAVIFT 437
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPFYEQYRGGQVSVCP 480
DB 438 QVSRQWRKRAEVTGTQIGSYFGFSLCSVDVDSNGSTDLVLIGPFPYEQYRGGQVSVCP 497
QY 481 PRQARWQCDAYLYGECQCPWCRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYL 540
DB 498 PRQWR-RWNCDAVLYGECQCPWCRFGAALTVLGDVNGDKLTDVVGAPGEENRGAYL 556
QY 541 HGTSGSGISPSHQRAGSLSPRLQYFGOSLSSGQDLTWDGLVDTLVGQGHVLLRSQ 600
DB 557 HGVLFPSISPSHQRAGSLSSRLQYFGQALSSGQDLTQDGLVLAAGARGQVLLRTR 616
QY 601 PVLRLVKAIMEFNEPREVARNVPECNDQVVKGEAGEVRVCLVHVKSTDRDRBQIQSVVT 660
DB 617 PVLWVGVMQFIPAEIPRSAFECEBQVVSQTLVQSNICLYIDKRSKNLIGSRDLQSSVT 676
QY 661 YDLALDSGRHSAVFNENSTRTQVGLTQCTETLKLQLPNCIEDPVSPTVLRNF 720
DB 677 LDALDPERLSPRATFOETKRSLSRVRLGLKAHCENFNLLPSCVEDSVPTILRNF 736
QY 721 SLVGTPLSAPCNLRPVLAEQALETALFFKXNCGNDNICQDDLSITFFPMSLDCILVWG 780
DB 737 TLVGKPELLAFNLRPMLAADAQRYETASLFFKXNCGADHICQDNLGISFSPPGLKSLV 796
QY 781 GPRFNVTVVRNDGDSYTVQTPFPPLDLSYRKVSTLQNRQSQRWRLACASSTEV 840
DB 797 SNLENAEVMVMNDGDSYCTTTFSPAGLSVRYVAEQKQQLRSLHLTCDSPAVG-- 854
QY 841 SGALKSTCSINHPPEPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPRTNTEF 900
DB 855 SQGTWSTSCINHLIFRGGQITFLATFDVSPRAVLGDRLLLTANVSSENTPRTSKTF 914
QY 901 QLELPKYAVVMVTSIGVSTKYLNTFAS-ENTSRVMQHQYOVGNLQORSILPISLVLVP 959
DB 915 QLELPKYAVVMVTSIGVSTKYLNTFAS-ENTSRVMQHQYOVGNLQORSILPISLVLVP 974
QY 960 VRLNQTVMDRPOVTFEENLSSTCTKRLPSHSDFLAELRKAPVNCSTAVCQRIQCDI 1019
DB 975 VELNQEAVMDVEVSLPQNSLSCSSSEKIAGPASDFLAHQKPNVLDSCSIAGCLFRCDV 1034
QY 1020 PFGIOEFENATLXGNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
DB 1035 PFSVQBELDFTLKGNLSPGVRQILQKYSVSVASITFDTSVYSQLPQGEAFWBAQT 1094
QY 1080 TKVPEFVFPNPLPIVGSVGGILLALLITAAALYKLGFEKQYKDMXSE 1128
DB 1095 TVLEKYVHNPTPLVIGSSIGGLLLALLITAVLYKVGFFKRYKEMMBE 1143
RESULT 13
AAR78166
ID AAR78166 standard; protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX
KW Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
XX inflammatory bowel disease; asthma.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 17..1108

FT Region /note= "extracellular domain"
 FT 150..352
 FT /note= "this region is homologous to the insert common to
 FT Chila,b,c and may be a site for interaction with ICAM
 FT family proteins"
 FT 465..474
 FT /note= "putative cation binding site"
 FT 518..527
 FT /note= "putative cation binding site"
 FT 592..600
 FT /note= "putative cation binding site"
 FT 1109..1128
 FT /note= "transmembrane region"
 FT 1129..1161
 FT /note= "cytoplasmic domain"
 PN W09517412-A1.
 PD 29-JUN-1995.
 XX 21-DEC-1994; 94WO-US014832.
 XX 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1995-240603/31.
 DR N-PSDB; AAQ91712.
 DR Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
 PT potential antiinflammatory agents, for the treatment of graft
 PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
 XX Claim 7; Page 82-87; 172pp; English.
 CC -Tm1 was based on a partial cDNA clone (given in AAQ91727) of canine alpha
 CC -Tm1 was used to screen a human spleen cDNA library to identify clone
 CC 1942 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO cells
 XX Sequence 1161 AA;
 SQ
 Query Match 58.2%; Score 3417; DB 2; Length 1161;
 Best Local Similarity 59.8%; Pred. No. 2.1e-275;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFQENARGFQGSVVLQGSRRVVVVGAPQEIIVAAVNAQGSGLYQCDYSTGSCBPI 60
 DB 17 FNLDVEEPTIFQEDAGGFGQSVVQFGGSLVVGAPLEVVAAVNAQGLYDCAANTGMCQPI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQILLACGPTVHQCSTENTVYKGLCFLPGSNLQCPQK 120
 DB 77 PLHIRPEAVNMSLGLTAAATNGSLRLLACGPTLHRCVCGNSYSKSGCLLGLSRW-EITQT 135
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMKBFVSTVMEQLKKSKTLFSLMOYSBEF 180
 DB 136 VPDATPECHQEMDIVFLIDGSGSIDQDNFQMGKPVQVMQFQESDTLPLALMQYSNLL 195
 QY 181 RHFTFPEKQNPFPRLVKPTTOLLGRTHATGVRKVTRELLNITNGARKAPKILLIVI 240
 DB 196 KIHFTPTQRTSPSQSLVDPIVLQKGLTFTATGILTVVTQLFHKNKGARKSAKILLIVI 255
 QY 241 TQGEKFGDPLGVEDVPIPEADRGVRYVIGVGDFAFRSEKSRQELNTIASKPPDRHVFQVN 300
 DB 256 TDGQKYKDPLEVDVPIPEAKAGIIEYALGVGHAFQGPARTARQELNTISSAPPQDHVFKVD 315
 QY 301 NEALATIQNQRKIPALTEGCTGSSSFEHMSQEGFSAITNGPLLSVTGSDYDWAG 360
 DB 316 NEAALGSIQKQKQKIYAVEGTQSRASSSFQHMSQEGFSTALTDGLFLGAVGFSWSG 375

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYGYAAAIILNRRVQSLVLAGPRYQHIGLVAWFR 420
 DB 376 GAFLYPNMSPPTFNNMQENVMDSDSYLGYSTELALWKGVQNLVLGAPRYQHTKAVIFT 435
 QY 421 QNTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSCPL 480
 DB 436 QVSRQWEKKAETGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSCPL 495
 QY 481 PRQORAEWQCDVLYGQGGQPGWGFGAALTVLGPNVNGDKLTDVAIGAPGEDNMGAYLYF 540
 DB 496 PRQORVQWQCDVLYRGQGGHGWGFGAALTVLGPNVEDKLDVAIGAPGEBQENRGAYLYF 555
 QY 541 HGTSGSISPSHSQRIAGSKLSRLQYFGSLSGGQDLTMDGLVDLTVGQKHVLLRSQ 600
 DB 556 HGASESISPSHSQRIASSQLSRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLRS 615
 QY 601 PVLKVKAIMBPNPREVARNVFECDNDQVVKGEAGEVVRVCLHVQKSTRDRLEBQIQSVVT 660
 DB 616 PVLKVGVMRFPSPVEVAKAVYRWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVAVNETKNSRRQTVLGLTQCTKLKQLPNCIEDPVSPIVRLNP 720
 DB 674 FDLALDPRLTSRAIPNETKNPTLTRRKTGLGTHCETLKLPLDCVEDVVSPIHLHNF 733
 QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNITCQDLSITFSPMSLDCLVVG 780
 DB 734 SLVREPISQNLRLPVLAVGSDLTFTASLPPEKNCQDGLCEGDLGYTLSPSGLQTLTVG 793
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLVYKRVSTLQNRQSRWRLACESASSTEV 840
 DB 794 SSLELVITVTVWVWAGDSYGTWVSLYFAGLSHRVSVGAQKQPHQSALRLACETV-PTED 852
 QY 841 SGALKSTSCSINHPIPPENSEVENTNITPDVDSKASLGNLKLKLLKANVTSENNMPTNKTET 900
 DB 853 EG-LRSSRCSVNHPIFHEGNGTFIVFDVSKATLDGRLMLRASSENKWKASSSKATP 911
 QY 901 QLELPVXYAVYVTVSHGVSTKYNL-TASENTSRVMOHQYQVSNLQGRSLPISLVLVLP 959
 DB 912 QLELPVXYAVYVTVSHGVSTKYNL-TASENTSRVMOHQYQVSNLQGRSLPISLVLVLP 971
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
 DB 972 VLLNGVAVWVWMEAPQSGL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029
 QY 1020 PFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1079
 DB 1030 PFSVQEEELDTLKNLSFGWVRETLQKVLVWVABITEDTSVYSQLPQGAFFVRSQTE 1089
 QY 1080 TKVFPFVPPNPLPLIVGSSVGLILLALITAAALYKLGFEKQYKDMKSE 1128
 DB 1090 MVLBEDEVYNAIPIIMGSSVGLILLALITATLYKLGFPFKHYKEMLED 1138
 RESULT 14
 AAW23049
 ID AAW23049 standard; protein; 1161 AA.
 XX AC AAW23049;
 XX DT 24-FEB-1998 (first entry)
 XX Human beta 2 integrin alpha d subunit.
 XX Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
 KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW psoriasis; lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 17..1108
 FT Domain

PT Peptide 1. .16
FT /note= "signal peptide"
PT Protein 17. .1161
FT /note= "mature protein"
PT Domain 17. .1108
FT /note= "extracellular domain"
FT Domain 150. .352
FT /note= "i (insertion) domain"
PT Region 1109. .1128
FT /note= "transmembrane region"
FT Domain 1129. .1161
FT /note= "cytoplasmic domain"
XX
XX US5766850-A.
PN
XX
XX 16-JUN-1998.
XX
PF 21-DEC-1994; 94US-00362652.
XX
XX 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
XX
XX (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
DR WPI; 1998-361678/31.
DR N-PSDB; AAV31540.
XX
XX Isolation of DNA encoding protein that binds to integrin subunit - using
PT recombinant cells containing reporter-transactivator construct.
XX
PS Example 5; Col 45-54; 86pp; English.
XX
XX This represents a human beta2 integrin alpha subunit (alpha d)
CC polypeptide. This is used in the methods of the invention for isolating a
CC polynucleotide encoding a protein that binds to alpha d. The method
CC comprises transforming or transfecting host cells with a DNA construct
CC comprising a reporter gene under the control of a promoter regulated by a
CC transcription factor having a DNA-binding domain and an activating
CC domain. A first hybrid DNA sequence encoding a fusion of at least part of
CC alpha d and either the DNA-binding domain or the activating domain of the
CC transcription factor and a library of hybrid DNA sequences encoding
CC fusions of at least part of putative alpha d-binding proteins and the DNA
CC -binding domain or the activating domain of the transcription factor
CC which is not incorporated in the first fusion are expressed in the host
CC cells. The binding of an alpha d-binding protein to alpha d in a
CC particular host cell is detected by determining production of the
CC reporter gene product in the cell. The hybrid DNA sequence encoding the
CC alpha d-binding protein can be isolated from the cell. Alpha d may be
CC useful for treating graft arteriosclerosis, atherosclerosis, diabetes,
CC inflammatory bowel disease, arthritis and multiple sclerosis
XX
SQ Sequence 1161 AA;
Query Match 58.2%; Score 3417; DB 2; Length 1161;
Best Local Similarity 59.8%; Pred. No. 2.1e-275;
Matches 675; Conservative 162; Mismatches 284; Indels 6; Gaps 6;
QY 1 FNLDENAMTFQENARFGQSVVLOGSRVWVCAPQEIIVAAHQSGSLXOCYSTGSCPI 60
DB 17 FNLDVEEPTIFQSDAGFGQSVVQFGSLRWVGAFLVVAARKTGILYDCAATGVCQFI 76
QY 61 RLQVPVAVNMSLGLSLAANTSPFOLLACGPTVHQTCSNTYVKGICFLFGNLRQPOQK 120
DB 77 PLHIREAVNMSLGLTLAASTNGSRULLACGPTLHRVCGENSYSGKSCLLGSRW-EIIQT 135
QY 121 FPALRGCPQEDSDIAFLIDGSGSIIPHPRKEFVSTVMEQLKSKTLFSLMOYSEEP 180
DB 136 VPDATPECPHQEMDIVFLIDGSGSIDQNDPNQMKGFVQAVMGQFEGDTLFLALMOYSNLL 195
QY 181 RIHFTKEFONNPNRSLVKPIQOLLGRHTATGVAKVIRELLNITNGARKNAFKILLIVI 240

Db 196 KIHFTTQFTSPSQSLVDPIVOLKGLFTTATGILTVTQTLFHHKNGARKSAKKILIVI 255
QY 241 TDGEKFGDPLGYEDVTPHADREGVIRYVIGVDAFRSEKSRQBLNTIASPPRDHVFQVN 300
Db 256 TDGQKYKDPLEYSDVIPAQKAGIIRYALGVGHAFQGPPTARQBLNTISSAPPQDHFVKVD 315
QY 301 NFEALKTIQNLREKIFAIEGTGTGSSSSPEHMSOBGFSAAITTSNGPLSTVGSYDWAG 360
Db 316 NFAALGSIQLOKEITYAVEGTOSRASSSQHMSQEGFTALTMGDLFLGAVOSFWSG 375
QY 361 GVFLYTSKEKSTFTNTRVDSMDNDAYLGYAAAIILNRRVQSLVGLGAPRYOHIGLVAMER 420
Db 376 GAFLYPNMGPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVGLGAPRYOHTKAVIFT 435
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILIGAPHYYEOTRGQGVSCPL 480
Db 436 QVSRQWKKAEVTTGTQIGSYFGASLCSVDVDSNGSTDLVILIGAPHYYEOTRGQGVSCPL 495
QY 481 PRGORARWQCDVLYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYLYF 540
Db 496 PRGORVWQCDVLYGRQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYLYF 555
QY 541 HGTSGGISPSHSHORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGRQGHVLLRSQ 600
Db 556 HGASEGISPSHSHORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGRQGHVLLRSQ 615
QY 601 PVLKVKVAMEFNPREEVARNVFECDQVVKGEAGEVVRVCLHVOKSTRDRLEGGQIQSVYT 660
Db 616 PVLKVKVAMEFNPREEVARNVFECDQVVKGEAGEVVRVCLHVOKSTRDRLEGGQIQSVYT 673
QY 661 YDLALDSGRPHSRVAPNETKNSRROTQVILGLTQCTETLKLQLPNCIEDPVPVILNLF 720
Db 674 YDLALDPGLRTSRAIFNETKNPTLTRKTLGLGTHCETLKLLEPDCVEDVVPVILNLF 733
QY 721 SLVCTPLSAFONRPLVLAEDAQRPLTALPPEKNCNDNICODDLSTIFSPMSLDCLVVG 780
Db 734 SLVREPSPONRPLVLAESQDLTASLPPEKNCNDNICODDLSTIFSPMSLDCLVVG 793
QY 781 GPRFNFVTVVRNDGDESYRTQVTFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 794 SSELNVITVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIPPENSEVTNITFEDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 853 EG-LRSSRCSYNHPIFHEGSGTFTVTDVSYKATLGDRLMMLRASSENKASSSKATP 911
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNF-TASENTSRVMQHOYQVSNLQORSPLSLVFLVP 959
Db 912 QLELPVKYAVVMVVTSHGVSTKYLNF-TASENTSRVMQHOYQVSNLQORSPLSLVFLVP 971
QY 960 VRLMGTVIWDRPQVTFPSENLSSTCHTERLPSSHDDFLAELKAPVWNCIAVCORIOCDI 1019
Db 972 VLLNGVAVWVWMEAPSQSL--PCVSEKRPQHSDFLTQISRSPLMDCSIADCLQFRCDV 1029
QY 1020 PFGIQEENFATLKNLSFDWVYIKTSHNHLIVSTABILFNDVSVFTLLPGQCAFVRSQTE 1079
Db 1030 PFSVQEBELDTLKNLSFGVWVRETLOKVLVSVABITFTDTSVVSQLPQGEAFMRAQWE 1089
QY 1080 TKVEFFEYFNPFLPLVGVSGGLLILALITAAALYKLGPFKQYKDMSE 1128
Db 1090 MVLEDEYVNAIPIIMGSSVAGALLALITATLYKLGPFKQYKDMSE 1138

Search completed: June 7, 2004, 17:11:58
Job time : 53.4608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds
(without alignments)
8225.189 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FRQYKQMSGCGPGRPEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5875	100.0	1137	10	US-09-902-481A-6
2	5868	99.9	1153	9	US-09-350-259-3
3	5868	99.9	1153	10	US-09-902-481A-1
4	5868	99.9	1153	10	US-09-891-943-3
5	5868	99.9	1153	14	US-10-144-259-30
6	5868	99.9	1153	14	US-10-207-655-176
7	5855	99.7	1137	10	US-09-902-481A-5
8	5852.5	99.6	1152	9	US-09-945-265-4
9	5845	99.5	1137	10	US-09-902-481A-3
10	5832	99.3	1137	10	US-09-902-481A-4
11	3473	59.1	1163	15	US-10-116-275-204
12	3450	58.7	1163	9	US-09-350-259-4
13	3450	58.7	1163	10	US-09-891-943-4
14	3417	58.2	1161	9	US-09-350-259-2
15	3417	58.2	1161	10	US-09-891-943-2

Sequence 99, Appl
Sequence 99, Appl
Sequence 55, Appl
Sequence 55, Appl
Sequence 53, Appl
Sequence 53, Appl
Sequence 37, Appl
Sequence 37, Appl
Sequence 46, Appl
Sequence 46, Appl
Sequence 1212, Ap
Sequence 1212, Ap
Sequence 1, Appl
Sequence 295, Appl
Sequence 1871, Ap
Sequence 1871, Ap
Sequence 103, App
Sequence 103, App
Sequence 101, App
Sequence 2, Appli
Sequence 2, Appli
Sequence 103, App
Sequence 103, App
Sequence 230, App
Sequence 810, App
Sequence 338, App
Sequence 35, Appl
Sequence 35, Appl
Sequence 4, Appli
Sequence 2, Appli

16 3401.5 57.9 1161 9 US-09-350-259-99
17 3401.5 57.9 1161 10 US-09-891-943-39
18 3239.5 55.1 1161 9 US-09-350-259-55
19 3239.5 55.1 1161 10 US-09-891-943-55
20 3228 54.9 1161 9 US-09-350-259-53
21 3228 54.9 1161 10 US-09-891-943-53
22 3218.5 54.8 1151 9 US-09-350-259-37
23 3218.5 54.8 1151 10 US-09-891-943-37
24 3207 54.6 1155 9 US-09-350-259-46
25 3207 54.6 1155 10 US-09-891-943-46
26 1848 31.5 369 12 US-10-087-192-1212
27 1552.5 26.4 1170 9 US-09-945-265-2
28 1548.5 26.4 1170 12 US-10-361-164-1
29 1525 26.0 1223 16 US-10-408-765A-295
30 1350.5 23.0 1086 16 US-10-408-765A-1871
31 1229.5 20.9 494 9 US-09-350-259-103
32 1229.5 20.9 494 10 US-09-891-943-103
33 1157.5 19.7 413 9 US-09-350-259-101
34 1157.5 19.7 413 10 US-09-891-943-101
35 1149 19.6 1179 14 US-10-177-550-2
36 1149 19.6 1179 15 US-10-173-551-2
37 1102.5 18.8 1151 10 US-09-894-130-103
38 1102.5 18.8 1151 10 US-09-836-353A-103
39 1102.5 18.8 1179 12 US-09-918-715-250
40 1101 18.7 1188 15 US-10-291-265-810
41 1097 18.7 1188 15 US-10-291-265-338
42 1093.5 18.6 1189 10 US-09-984-130-35
43 1093.5 18.6 1189 10 US-09-836-353A-35
44 1093.5 18.6 1189 12 US-10-262-839-4
45 1079 18.4 589 12 US-10-261-164-2

ALIGNMENTS

RESULT 1

US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 100.0%; Score 5875; DB 10; Length 1137;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPCEIVAANORGLVYOCDYSTGSCPT 60

1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPCEIVAANORGLVYOCDYSTGSCPT 60

61 RLQVPEAVNMSLGLSLAATTPPOLLAGCPVHTCSENIVYKGL FLFGNLRQPOK 120

61 RLQVPEAVNMSLGLSLAATTPPOLLAGCPVHTCSENIVYKGL FLFGNLRQPOK 120

121 FPEALRGCPQEDSDIAFLIDSGSIIIPDFRKNKEFVSTVMEQLKK KTLFSLMQISEEP 180

```

121 FPEALGCGCEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTVMEQLKSKTLFSLMOYSEBF 180
181 RHFTFKFQNNPNSLVKPIQLLGRTHATGVKVIKRELLNITNGARKNAFKILLIYI 240
181 RHFTFKFQNNPNSLVKPIQLLGRTHATGVKVIKRELLNITNGARKNAFKILLIYI 240
241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
301 NFPAKTIQNLREKIFAEGTGTSSSFEHMSQEGFSAATTSNGPLLSVTGSDWAG 360
301 NFPAKTIQNLREKIFAEGTGTSSSFEHMSQEGFSAATTSNGPLLSVTGSDWAG 360
361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHIGLVAMER 420
361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHIGLVAMER 420
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLI GAPHYVEOTRGQVSVCP 480
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLI GAPHYVEOTRGQVSVCP 480
481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDLTGAGQHVLIRSQ 600
541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDLTGAGQHVLIRSQ 600
601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720
721 SLVGTPLSAFNLRPVLAERDAORLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVG 780
721 SLVGTPLSAFNLRPVLAERDAORLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVG 780
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYKSTLQNRQSRSWRLACASSTEV 840
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYKSTLQNRQSRSWRLACASSTEV 840
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
901 QLELPVKYAVVWVTSKYSTKYLNTASNTSRVMOHQVQVSNLQORSLSPLSVFLVPV 960
901 QLELPVKYAVVWVTSKYSTKYLNTASNTSRVMOHQVQVSNLQORSLSPLSVFLVPV 960
961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKA PVVNCISIAVCORIQDIP 1020
961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDPLAELRKA PVVNCISIAVCORIQDIP 1020
1021 FFGIQEERNATLKNLSFDWVIKTSNHLITVSTAELIENDSVFTLLPGOGAFVRSOTET 1080
1021 FFGIQEERNATLKNLSFDWVIKTSNHLITVSTAELIENDSVFTLLPGOGAFVRSOTET 1080
1081 KVEPFEVNPPLIVGSSVGGILLALITAALYKLGFFKQYKDMSEGGPPGAEPP 1137
1081 KVEPFEVNPPLIVGSSVGGILLALITAALYKLGFFKQYKDMSEGGPPGAEPP 1137

```

RESULT 2
 US-09-350-259-3
 ; Sequence 3, Application US/09350259
 ; Patent No. US2002062008A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US2002062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,989
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

```

```

Query Match 99.9%; Score 5868; DB 9; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTPOENARGFGQSVVQIQGSRVVVVGAPQETVAANQORSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTPOENARGFGQSVVQIQGSRVVVVGAPQETVAANQORSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFPLFGNLRQPOPK 120
DB 77 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFPLFGNLRQPOPK 136
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTVMEQLKSKTLFSLMOYSEBF 180
DB 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTVMEQLKSKTLFSLMOYSEBF 196
QY 181 RHFTFKFQNNPNSLVKPIQLLGRTHATGVKVIKRELLNITNGARKNAFKILLIYI 240
DB 197 RHFTFKFQNNPNSLVKPIQLLGRTHATGVKVIKRELLNITNGARKNAFKILLIYI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNTIASKPRDHVFN 316
QY 301 NFPAKTIQNLREKIFAEGTGTSSSFEHMSQEGFSAATTSNGPLLSVTGSDWAG 360
DB 317 NFPAKTIQNLREKIFAEGTGTSSSFEHMSQEGFSAATTSNGPLLSVTGSDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHIGLVAMER 420
DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYOHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLI GAPHYVEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLI GAPHYVEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDLTGAGQHVLIRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSRLOVFGQSLGGQDLTMDGLVDLTGAGQHVLIRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
DB 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNTRRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRNLF 720

```

Db 677 YDLDSDGRPHSRVAFNETKSTTRQVGLGTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANKVTSENNMPTNKTEP 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANKVTSENNMPTNKTEP 916
QY 901 QLELPVKAVTVMTVSHGVSTKYLNFASNTSRVMQHOYQVSNLQGRSLPISLVPLVPV 960
Db 917 QLELPVKAVTVMTVSHGVSTKYLNFASNTSRVMQHOYQVSNLQGRSLPISLVPLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1036
QY 1021 PFGIOEBFNATLKGKLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1037 PFGIOEBFNATLKGKLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
QY 1081 KVEPEFVNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEFVNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 3

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RNK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIYVAANQSGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIYVAANQSGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVAVNNSIGLSLAATTPBOLLACGPTVHTQCSNTVYVKGCLFGLSNLRQCPQK 120
Db 77 RLQVPVAVNNSIGLSLAATTPBOLLACGPTVHTQCSNTVYVKGCLFGLSNLRQCPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQKSKTLFSLMOYSEEF 180

Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNPSLVKPIITQLGRTHATGVRKVTRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNPSLVKPIITQLGRTHATGVRKVTRELLNITNGARKNAFKILIVI 256
QY 241 TDCEKTCDDPLGYEDVTPEADREGVIRVVGDAFTSEKSRQELNITIASPPRDPHFQVN 300
Db 257 TDCEKTCDDPLGYEDVTPEADREGVIRVVGDAFTSEKSRQELNITIASPPRDPHFQVN 316
QY 301 NFBAKTIQNLAREKIPAIETGTGTGSSSSPEHEMSQEGFSAATITSGNPLLSITVGSYDWAG 360
Db 317 NFBAKTIQNLAREKIPAIETGTGTGSSSSPEHEMSQEGFSAATITSGNPLLSITVGSYDWAG 376
QY 361 GVFLYTSKEKSTINMTRVDSDAVDAYLGYAAAILLRNVQSLVLAGAPRYCHIGLVAMER 420
Db 377 GVFLYTSKEKSTINMTRVDSDAVDAYLGYAAAILLRNVQSLVLAGAPRYCHIGLVAMER 436
QY 421 QNTGWMESNANVKTQICAYFGASLSCVDVDSNGSTDVLVIGAPHYEYQTRGGQVSCPL 480
Db 437 QNTGWMESNANVKTQICAYFGASLSCVDVDSNGSTDVLVIGAPHYEYQTRGGQVSCPL 496
QY 481 PRGQARWQCDVLYGEGQGFWRPGALTVLDVNGDKLTDVAI GAPGEDNREGAVYLP 540
Db 497 PRGQARWQCDVLYGEGQGFWRPGALTVLDVNGDKLTDVAI GAPGEDNREGAVYLP 556
QY 541 HGTSGSGISPSHSORIASGKLSPLRYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKLSPLRYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMBPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRLEAGQIQSVVT 660
Db 617 PVLRVKAIMBPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRLEAGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTTRQVGLGTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTTRQVGLGTQTCETLKLQLPNCIEDPVSPIVLRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPFKKNCNDNICODDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANKVTSENNMPTNKTEP 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANKVTSENNMPTNKTEP 916
QY 901 QLELPVKAVTVMTVSHGVSTKYLNFASNTSRVMQHOYQVSNLQGRSLPISLVPLVPV 960
Db 917 QLELPVKAVTVMTVSHGVSTKYLNFASNTSRVMQHOYQVSNLQGRSLPISLVPLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDIP 1036
QY 1021 PFGIOEBFNATLKGKLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1037 PFGIOEBFNATLKGKLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
QY 1081 KVEPEFVNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEFVNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 4
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.

```

; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-3

Query Match          99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 76

QY 61 RLOQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGICFLFGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGICFLFGSNLRQPOK 136

QY 121 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEEF 180
DB 137 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEEF 196

QY 181 RHFTFKPONNPNRSLVKPIITOLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 240
DB 197 RHFTFKPONNPNRSLVKPIITOLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 256

QY 241 TDEKFGDPLGYEDVIPADREGVIRYVIGVDFAFRSEKSRQBLNTIASKPRDHVFQVN 300
DB 257 TDEKFGDPLGYEDVIPADREGVIRYVIGVDFAFRSEKSRQBLNTIASKPRDHVFQVN 316

QY 301 NFPAKTIQNLREKIPALBGTOTGSSSPHEHMSQEGSAATSNGLLSTVGSYDMAG 360
DB 317 NFPAKTIQNLREKIPALBGTOTGSSSPHEHMSQEGSAATSNGLLSTVGSYDMAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMESNANVKGITQIGVFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQVSCPL 480
DB 437 QNTGMESNANVKGITQIGVFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQVSCPL 496

QY 481 PRGQARWQCDVLYGSGQOPWGFPGAAALVLDVNGDKLTDVAIGAPEGEDNRGAVILF 540
DB 497 PRGQARWQCDVLYGSGQOPWGFPGAAALVLDVNGDKLTDVAIGAPEGEDNRGAVILF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLWDGLVLDLVTVGAQGVHLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDPLWDGLVLDLVTVGAQGVHLLRSQ 616

QY 601 PVLRVKALMFPNPREAVNPFECNDQVVKGEAGEVRVCHVOKSTDRLRQGIQSVVT 660
DB 617 PVLRVKALMFPNPREAVNPFECNDQVVKGEAGEVRVCHVOKSTDRLRQGIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 736

721 SLVCTPLSAFGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 780
737 SLVCTPLSAFGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 796

781 GPREPNVTTVVRNDGEDSYSTQVTFPEPLDI SVKXKUSTLONORSORSWRLACESASSTEV 840
797 GPREPNVTTVVRNDGEDSYSTQVTFPEPLDI SVKXKUSTLONORSORSWRLACESASSTEV 856

841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPTNKTPEP 900
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPTNKTPEP 916

901 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVSNL GORSLSISLVLVLPV 960
917 QLELPVKYAVTMVTVSHGVSTKYLNFNTASENTSRVMOHQVSNL GORSLSISLVLVLPV 976

961 RLNCTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCDIP 1020
977 RLNCTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCDIP 1036

1021 PFGIOBEFNATLKNLSFDWYIKTSHHLLIVSTABILFNDSVETLLPGOCAPVRSQDET 1080
1037 PFGIOBEFNATLKNLSFDWYIKTSHHLLIVSTABILFNDSVETLLPGOCAPVRSQDET 1096

1081 KVEPPEVNPPLIIVGSSVGGLLILALITAAALYKLGFFKRYQKDMMSGGPPGAEPPQ 1137
1097 KVEPPEVNPPLIIVGSSVGGLLILALITAAALYKLGFFKRYQKDMMSGGPPGAEPPQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arrasout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-259-30

Query Match          99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 76

QY 61 RLOQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGICFLFGSNLRQPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGICFLFGSNLRQPOK 136

QY 121 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEEF 180
DB 137 FPEALGCGQEDSDIAFLIDGSGSIIPHDFRMRKPEFVSTMQLKSKTFLSLMOYSEEF 196

QY 181 RHFTFKPONNPNRSLVKPIITOLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 240
DB 197 RHFTFKPONNPNRSLVKPIITOLGRTHGTATGKVRKRELLNITNGARKNAFKLLIYI 256
```



```
QY 241 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPRDHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTSSTSSFFHEMSQEGFSAITNSGELLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTSSTSSFFHEMSQEGFSAITNSGELLSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWSNANVKGQIAGYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQVSVCP 480
DB 437 QNTGWSNANVKGQIAGYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEGQPMGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEGQPMGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKSTRQTVGLGTQTCETLKLQPCIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRAVFNETKSTRQTVGLGTQTCETLKLQPCIEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFGMLRPVLAEDAORLTALPFPFKKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGMLRPVLAEDAORLTALPFPFKKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTPFPFPLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTPFPFPLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 856
QY 841 SGALKSTCSINHPFPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTCSINHPFPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNFASNTSRVMQHOYQVSNLQORSILPISLVLV 960
DB 917 QLELPVKYAVVMVTSHGVSSTKYLNFASNTSRVMQHOYQVSNLQORSILPISLVLV 976
QY 961 RLNQTVIWRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVVMNCSTAVCORIQCDIP 1020
DB 977 RLNQTVIWRPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVVMNCSTAVCORIQCDIP 1036
QY 1021 FFGIQEENATLKENLSFQWYIKTSENHLLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1080
DB 1037 FFGIQEENATLKENLSFQWYIKTSENHLLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1096
QY 1081 KVEPFVPMPLPIVGVSSVGLLLALLITAAALYKLGFFKQKDMMSBGGPGAEPQ 1137
DB 1097 KVEPFVPMPLPIVGVSSVGLLLALLITAAALYKLGFFKQKDMMSBGGPGAEPQ 1153
```

RESULT 6

```
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
```

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 176

; LENGTH: 1153

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-207-655-176

Query Match 99.9%; Score 5868; DB 14; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFQENARFGQSVVQLOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPPVAVANVSGLSLAATTPPOLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 120
DB 77 RLQVPPVAVANVSGLSLAATTPPOLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRRMEKFVSTVMEQKKSKTLPSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRRMEKFVSTVMEQKKSKTLPSLMQYSEEF 196
QY 181 RIHTFPKQFQNNPRSLVVKPIITOLGRTHATGVRVIRELLNITNGARKNAFKILLVI 240
DB 197 RIHTFPKQFQNNPRSLVVKPIITOLGRTHATGVRVIRELLNITNGARKNAFKILLVI 256
QY 241 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVIVPEADREGVIRVYGVGDAPFRSEKSRQELNTIASKPRDHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTSSTSSFFHEMSQEGFSAITNSGELLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTSSTSSFFHEMSQEGFSAITNSGELLSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWSNANVKGQIAGYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQVSVCP 480
DB 437 QNTGWSNANVKGQIAGYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEGQPMGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEGQPMGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKSTRQTVGLGTQTCETLKLQPCIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRAVFNETKSTRQTVGLGTQTCETLKLQPCIEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFGMLRPVLAEDAORLTALPFPFKKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGMLRPVLAEDAORLTALPFPFKKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTPFPFPLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTPFPFPLDLSYRKVSTLQNRQSRQSWPLACESASSTEV 856
QY 841 SGALKSTCSINHPFPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTCSINHPFPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNFASNTSRVMQHOYQVSNLQORSILPISLVLV 960
```

Db 917 QLELPVYAYVMVTSCHVSKYLNFTASNTSRVMOHQVQVSNLGRSLPISLVLFPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCORIOCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCORIOCDIP 1036
Qy 1021 FPGIOEFNATLKNLSFDMVYKTSNHLHLLVSTABILLFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FPGIOEFNATLKNLSFDMVYKTSNHLHLLVSTABILLFNDVSFTLLPGQAFVRSQTET 1096
Qy 1081 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSBEGPPGAEPQ 1137
Db 1097 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSBEGPPGAEPQ 1153

RESULT 7
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 99.7%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGQGVVQLQGSRRVVGAPQBIIVANQSGSYQCDYSTGSCPI 60
Db 1 FNLDTENAMTQENARGQGVVQLQGSRRVVGAPQBIIVANQSGSYQCDYSTGSCPI 60

Qy 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVKGCLCFPLGSLNLRQOPQK 120
Db 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVKGCLCFPLGSLNLRQOPQK 120

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKXKTLFSLMQYSEEP 180
Db 121 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKXKTLFSLMQYSEEP 180

Qy 181 RIHFTKEFQNNPRSLVKPIITQLGKTHATGVRKVIKRELLTNGARKNAFKILIVI 240
Db 181 RIHFTKEFQNNPRSLVKPIITQLGKTHATGVRKVIKRELLTNGARKNAFKILILI 240

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDFVQVN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDFVQVN 300

Qy 301 NPEALKTONQRLREKIPAIETQGTSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAQ 360
Db 301 NPEALKTONQRLREKIPAIETQGTSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAQ 360

Qy 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAILNRVQSLVIGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAILNRVQSLVIGAPRYOHIGLVAMFR 420

Qy 421 QNTGWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGGQSVUCPL 480

Db 421 QNTGWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGGQSVUCPL 480
Qy 481 PRGQBARWQCCDAVLYGEOQPMGRGAALTVLGDVNGDKLTDVAIGAPEEDNRGAVLYP 540
Db 481 PRGQBARWQCCDAVLYGEOQPMGRGAALTVLGDVNGDKLTDVAIGAPEEDNRGAVLYP 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSGSLSGQDLTMDGLVLTIVGAQGHVILLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSGSLSGQDLTMDGLVLTIVGAQGHVILLRSQ 600
Qy 601 PVLRYKALMEFNPREVARNVFECDNVVKGKAGVRVCLHVQKSTRDLRSGQIQSVVT 660
Db 601 PVLRYKALMEFNPREVARNVFECDNVVKGKAGVRVCLHVQKSTRDLRSGQIQSVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKSTRQTQVIGELTOTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 661 YDLALDSGRPHSRAVFNETKSTRQTQVIGELTOTCETLKLQLPNCIEDPVSPVILRLNF 720
Qy 721 SLVGTPLSAPGNLRPVLAEDAORLFTALPPFKKNGNDNICODELSITFESFMSLDCLVVG 780
Db 721 SLVGTPLSAPGNLRPVLAEDAORLFTALPPFKKNGNDNICODELSITFESFMSLDCLVVG 780
Qy 781 GRBENVTVVRNDGEDSVRTQVTPFFPLDLSYRKVSTLQNSQSRSWELACESASTEV 840
Db 781 GRBENVTVVRNDGEDSVRTQVTPFFPLDLSYRKVSTLQNSQSRSWELACESASTEV 840
Qy 841 SGALKSTCSINHPITPPENSEVTENITFDVDSKASLGNKLLKANTSENMPRTNKTEF 900
Db 841 SGALKSTCSINHPITPPENSEVTENITFDVDSKASLGNKLLKANTSENMPRTNKTEF 900
Qy 901 QLELPVYAYVMVTSCHVSKYLNFTASNTSRVMOHQVQVSNLGRSLPISLVLFPV 960
Db 901 QLELPVYAYVMVTSCHVSKYLNFTASNTSRVMOHQVQVSNLGRSLPISLVLFPV 960
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCORIOCDIP 1020
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCORIOCDIP 1020

Qy 1021 FPGIOEFNATLKNLSFDMVYKTSNHLHLLVSTABILLFNDVSFTLLPGQAFVRSQTET 1080
Db 1021 FPGIOEFNATLKNLSFDMVYKTSNHLHLLVSTABILLFNDVSFTLLPGQAFVRSQTET 1080

Qy 1081 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSBEGPPGAEPQ 1137
Db 1081 KVEPPEVNPPLPIVGVSSVGGILLALITAALYKLGFFKQYKDMMSBEGPPGAEPQ 1137

RESULT 8
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimada, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CEN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.6%; Score 5852.5; DB 9; Length 1152;
Best Local Similarity 99.6%; Pred. No. 0;

Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;			
Qy	1	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPQEI	VAANORGSLYQCDYSTGSCPEI 60
Db	17	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPQEI	VAANORGSLYQCDYSTGSCPEI 76
Qy	61	RLOVPVAVNMVSLGLSLAATTSPPQILACGPTVHQC	SENTYVKGICFLFGSNLRQOPQK 120
Db	77	RLOVPVAVNMVSLGLSLAATTSPPQILACGPTVHQC	SENTYVKGICFLFGSNLRQOPQK 136
Qy	121	FPFALRGCPQEDSDIAFLDGSQSIIPHDPRRMKBFV	STMVMEQKSKTLFSLMQYSEEF 180
Db	137	FPFALRGCPQEDSDIAFLDGSQSIIPHDPRRMKBFV	STMVMEQKSKTLFSLMQYSEEF 196
Qy	181	RIHFTFKFQNNPNRSLVPIIQLLGRTHATGVRKVI	RELLNITNGARKNAFKILIVI 240
Db	197	RIHFTFKFQNNPNRSLVPIIQLLGRTHATGVRKVI	RELLNITNGARKNAFKILIVI 256
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSE	KSRQBLNTIASKPPRDHVPQVN 300
Db	257	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSE	KSRQBLNTIASKPPRDHVPQVN 316
Qy	301	NFEALKTIQNLREKIPAIETGTOTGSSSSPEHEMS	QEGFSAAITNSGPIILSTVGSYDNAG 360
Db	317	NFEALKTIQNLREKIPAIETGTOTGSSSSPEHEMS	QEGFSAAITNSGPIILSTVGSYDNAG 376
Qy	361	GVFLYTSKSKSTFNNMTRVDSMDNDAYLGVAALII	LRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db	377	GVFLYTSKSKSTFNNMTRVDSMDNDAYLGVAALII	LRNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy	421	QNTGWSNSANVKTQIGVFGASLCSVDVDSNGSTD	LVLIIGAPHYVEOTRGQSVCP 480
Db	437	QNTGWSNSANVKTQIGVFGASLCSVDVDSNGSTD	LVLIIGAPHYVEOTRGQSVCP 496
Qy	481	PRGQARWQCDVAVLYGEGQGPWGRFGAALTVDG	VNGDKLTVDVAIGAPBEDNRGAVYLF 540
Db	497	PRGQARWQCDVAVLYGEGQGPWGRFGAALTVDG	VNGDKLTVDVAIGAPBEDNRGAVYLF 555
Qy	541	HGTSGSGISPSHSORIASGSKSLPQYFGQSLSGQ	DLTMDGLVDLTGCAQGHVLLRSQ 600
Db	556	HGTSGSGISPSHSORIASGSKSLPQYFGQSLSGQ	DLTMDGLVDLTGCAQGHVLLRSQ 615
Qy	601	PVLAVKAIMFNPREVARNVFCNDQVVKGEAGEVR	VLHVOKSTRDLREGQIQSVVT 660
Db	616	PVLAVKAIMFNPREVARNVFCNDQVVKGEAGEVR	VLHVOKSTRDLREGQIQSVVT 675
Qy	661	YDLALDSGRPHSRVAVNETKSTRQTVGLGTQCTE	TKLQPNICIEDPVPVILRLNF 720
Db	676	YDLALDSGRPHSRVAVNETKSTRQTVGLGTQCTE	TKLQPNICIEDPVPVILRLNF 735
Qy	721	SLVGTPLSAFGLNRPVLAEDAQRLEFTALFPFEK	YCGNDNICQDDLSITFSFMSLDCLVVG 780
Db	736	SLVGTPLSAFGLNRPVLAEDAQRLEFTALFPFEK	YCGNDNICQDDLSITFSFMSLDCLVVG 795
Qy	781	GPRFNTVTVVRNDGDSYRTQVTFPPPLDLVSRK	VSTLQNRQSRWRLACESASTEV 840
Db	796	GPRFNTVTVVRNDGDSYRTQVTFPPPLDLVSRK	VSTLQNRQSRWRLACESASTEV 855
Qy	841	SGALKSTSCSINRIPFENSEVFNITFDVDSKAS	LNKLLKXANTYSENMPFNKTEF 900
Db	856	SGALKSTSCSINRIPFENSEVFNITFDVDSKAS	LNKLLKXANTYSENMPFNKTEF 915
Qy	901	QLELPVKYAVVMVTSKGVTKYLNFTASENTSRV	MOHVOYNSLQORSLPISLVLFPV 960
Db	916	QLELPVKYAVVMVTSKGVTKYLNFTASENTSRV	MOHVOYNSLQORSLPISLVLFPV 975
Qy	961	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVNVNCSIAVCQIQCDIP 1020
Db	976	RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVNVNCSIAVCQIQCDIP 1035
Qy	1021	PFGLQEBFNATLKNLSFDVAVIKTSNNHLLIVST	AEILNDSVFTLLRQGGPVSQOTEF 1080
Db	1036	PFGLQEBFNATLKNLSFDVAVIKTSNNHLLIVST	AEILNDSVFTLLRQGGPVSQOTEF 1095

RESULT 9
 US-09-902-481a-4
 ; Sequence 4, Application US/09902481A
 ; Publication No. US20030054440A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Springer, Timothy
 ; APPLICANT: Shimaoka, Motomu
 ; APPLICANT: Shifman, Julia
 ; APPLICANT: Mayo, Stephen
 ; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
 ; FILE REFERENCE: A-70586-1/REF/RMS/RMK
 ; CURRENT APPLICATION NUMBER: US/09/902,481A
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/216,600
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1137
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic
 US-09-902-481a-4

Query Match 99.5%; Score 5845; DB 10; Length 1137;			
Best Local Similarity 99.0%; Pred. No. 0;			
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;			
Qy	1	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPQEI	VAANORGSLYQCDYSTGSCPEI 60
Db	1	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPQEI	VAANORGSLYQCDYSTGSCPEI 60
Qy	61	RLOVPVAVNMVSLGLSLAATTSPPQILACGPTVHQC	SENTYVKGICFLFGSNLRQOPQK 120
Db	61	RLOVPVAVNMVSLGLSLAATTSPPQILACGPTVHQC	SENTYVKGICFLFGSNLRQOPQK 120
Qy	121	FPFALRGCPQEDSDIAFLDGSQSIIPHDPRRMKBFV	STMVMEQKSKTLFSLMQYSEEF 180
Db	121	FPFALRGCPQEDSDIAFLDGSQSIIPHDPRRMKBFV	STMVMEQKSKTLFSLMQYSEEF 180
Qy	181	RIHFTFKFQNNPNRSLVPIIQLLGRTHATGVRKVI	RELLNITNGARKNAFKILIVI 240
Db	181	RIHFTFKFQNNPNRSLVPIIQLLGRTHATGVRKVI	RELLNITNGARKNAFKILIVI 240
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSE	KSRQBLNTIASKPPRDHVPQVN 300
Db	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAFRSE	KSRQBLNTIASKPPRDHVPQVN 300
Qy	301	NFEALKTIQNLREKIPAIETGTOTGSSSSPEHEMS	QEGFSAAITNSGPIILSTVGSYDNAG 360
Db	301	NFEALKTIQNLREKIPAIETGTOTGSSSSPEHEMS	QEGFSAAITNSGPIILSTVGSYDNAG 360
Qy	361	GVFLYTSKSKSTFNNMTRVDSMDNDAYLGVAALII	LRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db	361	GVFLYTSKSKSTFNNMTRVDSMDNDAYLGVAALII	LRNRVQSLVGLGAPRYQHIGLVAMFR 420
Qy	421	QNTGWSNSANVKTQIGVFGASLCSVDVDSNGSTD	LVLIIGAPHYVEOTRGQSVCP 480
Db	421	QNTGWSNSANVKTQIGVFGASLCSVDVDSNGSTD	LVLIIGAPHYVEOTRGQSVCP 480
Qy	481	PRGQARWQCDVAVLYGEGQGPWGRFGAALTVDG	VNGDKLTVDVAIGAPBEDNRGAVYLF 540
Db	481	PRGQARWQCDVAVLYGEGQGPWGRFGAALTVDG	VNGDKLTVDVAIGAPBEDNRGAVYLF 540
Qy	541	HGTSGSGISPSHSORIASGSKSLPQYFGQSLSGQ	DLTMDGLVDLTGCAQGHVLLRSQ 600

DB 541 HGTSGSGISPSHQSRIAGSKLSPRLOYRQSGSLSGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
QY 601 PVLRYKAIEMFNPVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRYKAIEMFNPVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVFNETHKSTRQTCVGLGTTCCETLKLQPNCEIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRVFNETHKSTRQTCVGLGTTCCETLKLQPNCEIEDPVSPIVLRNF 720
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPFEKNCNDNICODDLSITPSFMSLDCLVVG 780
DB 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPFEKNCNDNICODDLSITPSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQORSORSWRLACASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQORSORSWRLACASSTEV 840
QY 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEP 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEP 900
QY 901 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLVFLVPV 960
DB 901 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLVFLVPV 960
QY 961 RLNQTVIWDREQVTFSENLSTCHTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
DB 961 RLNQTVIWDREQVTFSENLSTCHTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
QY 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAVRSQTET 1080
DB 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAVRSQTET 1080
QY 1081 KVEPEFVNPLPIVGVSSVGLLLALLITAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1081 KVEPEFVNPLPIVGVSSVGLLLALLITAALYKLGFFKQYKDMMSGGPPGABEQ 1137

RESULT 10

US-09-902-481a-3
; Sequence 3, Application US/09902481a
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481a
; PRIOR FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TIPS: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481a-3

Query March 99.3%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVLOGSRVVGAPQBIIVAAHQSLYQCDYSTGSCETP 60
DB 1 FNLDTENAMTQENARGFGQSVVLOGSRVVGAPQBIIVAAHQSLYQCDYSTGSCETP 60
QY 61 RLQVPEAVNNSLGLSLAATTSPPQLLACGPTVHTQTSENTYVYKGLCFPLFGSNLRQOPQK 120

RESULT 11

DB 61 RLQVPEAVNNSLGLSLAATTSPPQLLACGPTVHTQTSENTYVYKGLCFPLFGSNLRQOPQK 120
QY 121 PPEALRGCPQSDSDIAPLIDGSGSIIIPHDPRMKPEFYSTVMBOLKSKTLPSLMOYSEEF 180
DB 121 PPEALRGCPQSDSDIAPLIDGSGSIIIPHDPRMKPEFYSTVMBOLKSKTLPSLMOYSEEF 180
QY 181 RIHFTFKFQFQNPNSRSLIKPITQLGRTHATGLRKVRRELFTNITNGARKNAFKILFLL 240
DB 181 RIHFTFKFQFQNPNSRSLIKPITQLGRTHATGLRKVRRELFTNITNGARKNAFKILFLL 240
QY 241 TDGKFKGDPGLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNLTASKPPRDHVPQVN 300
DB 241 TDGKFKGDPGLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNLTASKPPRDHVPQVN 300
QY 301 NPEALKTIQNLQREKI PAIBEGTQTSSSSEHMSQEGFSAATSTGSPILLSTVGSYDWAG 360
DB 301 NPEALKTIQNLQREKI PAIBEGTQTSSSSEHMSQEGFSAATSTGSPILLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFNNMTRVDSMDNDAYLGAAAIILNRRVQSLVLAGAPRYOHIGLVAMPR 420
DB 361 GVFLYTSKEKSTFNNMTRVDSMDNDAYLGAAAIILNRRVQSLVLAGAPRYOHIGLVAMPR 420
QY 421 QNTGMBESNANVKCTOIGAYEGASLCSVDVDSNGSTDLVILGAPHYEOTRGQGVSCPL 480
DB 421 QNTGMBESNANVKCTOIGAYEGASLCSVDVDSNGSTDLVILGAPHYEOTRGQGVSCPL 480
QY 481 PRGQARWQCDVLYGSGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540
DB 481 PRGQARWQCDVLYGSGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540
QY 541 HGTSGSGISPSHQSRIAGSKLSPELOYFGSLSGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHQSRIAGSKLSPELOYFGSLSGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
QY 601 PVLRYKAIEMFNPVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRYKAIEMFNPVARNVPECNDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVFNETHKSTRQTCVGLGTTCCETLKLQPNCEIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRVFNETHKSTRQTCVGLGTTCCETLKLQPNCEIEDPVSPIVLRNF 720
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPFEKNCNDNICODDLSITPSFMSLDCLVVG 780
DB 721 SLVGTPLSAFGLNRPVLAEDAQRFTALPPFEKNCNDNICODDLSITPSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQORSORSWRLACASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQORSORSWRLACASSTEV 840
QY 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEP 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGKLLKANVTSENNMPTNKTEP 900
QY 901 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLVFLVPV 960
DB 901 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLVFLVPV 960
QY 961 RLNQTVIWDREQVTFSENLSTCHTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
DB 961 RLNQTVIWDREQVTFSENLSTCHTKERLPSHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
QY 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAVRSQTET 1080
DB 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAVRSQTET 1080
QY 1081 KVEPEFVNPLPIVGVSSVGLLLALLITAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1081 KVEPEFVNPLPIVGVSSVGLLLALLITAALYKLGFFKQYKDMMSGGPPGABEQ 1137

138 LFVSRQCEPRQODIVFLIDSGSSISSENFATMNFVRAVISQORSTQSLMQFENKF 197
181 RHFTFKPQNNPNPSLVKPTQTLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
198 QTHFTFEPRSTNSPLSLASVHQLQGTFTYATATQNVVHRLPHASYGARDAIKILIVI 257
241 TDGKFGDPLGYEDVPEADREGVRYVIGVDAPRSEKSRQELNNTIASKPPRDHVFQVN 300
258 TDGKKGSDLDYKQVWPADAAGIIRYAIYVGLAFQNRNSWKLNDIASKPSQEHIFKVE 317
301 NFEALKTIONQRLREKIPALEGTQSGSSSEPHMSQGFSAATISNGPILLSVGSYDAG 360
318 DFDALKQIQNLQKKEKIPALEGTETISSSTFELENAQGFSAVTPDGPVLGAVGSPFWSG 377
361 GYFLYTSKESKSTFINMTVRDSDMDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
378 GAFLYPPNNSPTFINNSQENVMDRDSYLGSTELALWKGVQSLVGLGAPRYOHIGKAVIFI 437
421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCPPL 480
438 QVSRQWRMKAQVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYETQTRGGQSVCPPL 497
481 PRGQARWCDDAVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
498 PRGWR-RWVCDVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
541 HGTSGSISPSHRSORIASKSLSPRIQYQGSISGGQDLTMDGLVDLTGVAGQHVLLRSQ 600
557 HGVLGSPISPSHRSORIASQSLSSRLQYFGALSQGGQDLTQDGLVDLAVGARGOVLLRTR 616
601 PVLRYKAIEMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRRLRSGQIQSVVT 660
617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICLVYIDKRSKNLLGSRDLQSSVT 676
661 YDLALDSCRPSRAVFNETKNSRQTVLGLTCTCTETLKLQLPNCIEDPVSPIVLRINF 720
677 LDALAPGLSPRAIFQETKRSLSRVKVLGKAKCENFNLLPSCEVDSVPIILRLNF 736
721 SLVGTPLGAFGNLRPLAEDAQRLFTALPPEKKNCGMDNICODDLSITPSFMSLDCLVWG 780
737 TLVKGFLAFLNRLPLAALAQRYFTASLPPEKKNCGADHICQDNLGIFSFPGLKSLVG 796
781 GPRENVTVVRNDEGDSYRQVTPFPPLDLSYKSVSTLQNRQSRQWRLACSSASTEV 840
797 SNLELNARVWVMDGSDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
841 SGALKSTSCSINHPIFPENSBVTENITFDVDSKASLGNKLLKANVTSENNMPTKTEF 900
855 SGTWSTSCRINHILFRGAQITFLATFDVSPKAVGLDRULLLIANYSENNIERTSKTIF 914
901 QLELPVKYAVYVWVTSYHGVSTKYLNPTAS-BNTRVMOHQVQVSNLQORSILPISLFLVP 959
915 QLELPVKYAVYVWVSSHEQFTKYLNFSSEKESHVAMERYQVNNLQORDLPYSINFWP 974
960 VRLNQTIVDRPQVTFSENLSSTCTKRLPSSHDFLAELRKAPVNVNCSTAVCORIQCDI 1019
975 VELNQAQVMDVEVSHPCNPISLSCSSKIPAPASDPLAHIQKVPVLDCSIAGCLRFECDV 1034
1020 PFFGLOEFNATLKNLSFDWYKTSSENHLLIVSTABILFNDSVFTLLPQCGAFVRSQTE 1079
1035 PSFSVQEEELDFLKNLSFGVVRQIQKQVSNVSAEIIFTSVISQLPQGEAFMRAQTI 1094
1080 TKVPPFVPEPLIVGVSSVGLLALLALITAALYKLGFFKRVQKQVMDMSE 1128
1095 TVLEKYKVENPILIVGVSSIGLLALLALITAVLYKVGFFKRVQKEMWEE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007729A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007729A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-15
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.7%; Score 3450; DB 10; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7.4e-308;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGPQSVWOLQSGRVVVGAPQEIIVANQSGSLYQCDYSTGSCBPI 60
DB 20 FNLDTEELTAFRVDSAGFGDSVQVYANSVWVGAPQKIIAANOIGLYQCGYSTGACBPI 79
QY 61 RLQVPEAVNNSLGLSLAATTPQLLACGPTVHTQTCENTYVKGCLFLFQSNLRQOPQK 120
DB 80 GLQVPEAVNNSLGLSLASTTSPQLACGPTVHCEGRNMYLGLCFLGPT--QLTOR 137
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIHDFRMEKEFVSTVMEQKSKTLFSLMOYSEEP 180
DB 138 LFVSRQCEPRQODIVFLIDSGSSISSENFATMNFVRAVISQORSTQSLMQFENKF 197
QY 181 RHFTFKPQNNPNPSLVKPTQTLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
DB 198 QTHFTFEPRSTNSPLSLASVHQLQGTFTYATATQNVVHRLPHASYGARDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVPEADREGVRYVIGVDAPRSEKSRQELNNTIASKPPRDHVFQVN 300
DB 258 TDGKKGSDLDYKQVWPADAAGIIRYAIYVGLAFQNRNSWKLNDIASKPSQEHIFKVE 317
QY 301 NFEALKTIONQRLREKIPALEGTQSGSSSEPHMSQGFSAATISNGPILLSVGSYDAG 360
DB 318 DFDALKQIQNLQKKEKIPALEGTETISSSTFELENAQGFSAVTPDGPVLGAVGSPFWSG 377
QY 361 GYFLYTSKESKSTFINMTVRDSDMDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 378 GAFLYPPNNSPTFINNSQENVMDRDSYLGSTELALWKGVQSLVGLGAPRYOHIGKAVIFI 437
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCPPL 480
DB 438 QVSRQWRMKAQVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYETQTRGGQSVCPPL 497
QY 481 PRGQARWCDDAVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 498 PRGWR-RWVCDVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
QY 541 HGTSGSISPSHRSORIASKSLSPRIQYQGSISGGQDLTMDGLVDLTGVAGQHVLLRSQ 600
DB 557 HGVLGSPISPSHRSORIASQSLSSRLQYFGALSQGGQDLTQDGLVDLAVGARGOVLLRTR 616
QY 601 PVLRYKAIEMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRRLRSGQIQSVVT 660
DB 617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICLVYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSCRPSRAVFNETKNSRQTVLGLTCTCTETLKLQLPNCIEDPVSPIVLRINF 720
DB 677 LDALAPGLSPRAIFQETKRSLSRVKVLGKAKCENFNLLPSCEVDSVPIILRLNF 736

APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US20030077278A1e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1161
TYPE: PRT
ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.2%; Score 3417; DB 10; Length 1161;
Best Local Similarity 59.8%; Pred. No. 8.1e-305;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTFOENARFGOSVVOLOGSRVVGAPQEIYAANORGSLYQCDYSTGSCEPI 60
Db 17 FNLDVEEPTTFQEDAGFGOSVVOFGSRVVGAPLEVAANTGRVYDCAAIGHCQPI 76

Qy 61 RLQVPEAVNMSGLSLAATSPQLLACGFTVHQTCSNTYKGLCFLFGSNLRQOPQK 120
Db 77 PLHIRPEAVNMSGLTAASTNGSRLLACGFTLHRVCGENSYSKSGCLLIGSRW-BIIQT 135

Qy 121 FPEALRCPOEDSDIAFLIDGSGIIPHDPRMKEFYSTWMEQLKSKTLFSLMOYSEEP 180
Db 136 VPDATPECPHQEMDIVFLIDGSGIDQNDENQMKGFVQAVNGQFEGFTLFLALMOYSNLL 195

Qy 181 RIHFTFEFQNNPRSLVXPIITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
Db 196 KIHFTFTQFRTSPSQSLVDPVQLKGLTFTATGILTVVTQLFHHKNGARKSAKKILIVI 255

Qy 241 TDGKFGDPLGYEDVPEADREGVIRVIGVDAPFRSEKSRQELNITIASKPRDHVQVN 300
Db 256 TDGQKYDPLYSVDVPOAEKAGIIRYAIGVGHAFQGPARTAOELNITISSAPPQDHVFKVD 315

Qy 301 NFEALKTONLREKI FALGETQTGSSSSFEHENSQGFSAATSNGLPILLSVGSYDWAQ 360
Db 316 NFAALGSIQXQLQEKIYAVEGTQSRASSPQFHEMSQEGFSTALTDGLFLCANGSPSBSG 375

Qy 361 GVFLYTSKEKSTFINMTRVDSMDKDAVILGYAAAIIILNRVQSULVIGAPRYOHIGLVAMFR 420
Db 376 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVILGAPRYOHTGKAVIFT 435

Qy 421 QNTGWESNANVKTGTOIGAYFGASLCSYDVDSNGSTDLVILGAPHYVQTRGGQVSVCP 480
Db 436 QVSRQWRKKAFTVGTQIGSYFGASLCSYDVDSNGSTDLVILGAPHYVQTRGGQVSVCP 495

Qy 481 PRGQARWQCDAVILGEGCQFPGWFGAALTVLGVNGDKLTDAIGAPEGEDNRGAVYLF 540
Db 496 PRGQVQVQCDAVILGEGCQFPGWFGAALTVLGVNEDKLDVAIGAPEGQENRGAVYLF 555

Qy 541 HGTSGSGISPHSGQRIAGSKLSPLQYFGOSLSCGODLTMDGLVDLTVCAGHVILLRSQ 600
Db 556 HGASESGISPHSGQRIAGSKLSPLQYFGQALSGGQDLTDGLMDLAVGARGQVILLRSL 615

Qy 601 PVLRYKALMEFNPREVARNVECNQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLKVGVMRSPVEVAAYKCEKEPSALEAGDATVCLTIQKSSLDL--GDIQSSVR 673

Qy 661 YDLALDSQPHSRVAFNETKNSRTROTQVILGLTQTCETLKQLPNCIEDPVPVILNLF 720
Db 674 PDLALDPGLRTSRAIFNETKNTPLTERRKTLGLIGHCETLKLLLPDCVEDVVPVILNLF 733

721 SLVGTPLSAFGNLRPLVLAEDAQRULTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 780
734 SLVREPIPSQNLRLPVLAVGSDLTASLPFKKNCQDGLCEGDLGVTLSPGLQTLTVG 793
781 GPREPNVTIVRNDGEDSYRTQVTFFPFLDLSYKRVSTLQNRQSRSWRLACESASSTEV 840
794 SSLELNVITVWNAGEDSYGTVVSLYFAGLSHRVSGAQKQPHQSALRLACETV-PTED 852
841 SGALKSTCSNHPITFPENSEVTNITFDVDSKASLGNKLLKAVTSENMPRTNKTEP 900
853 EG-LSSRCVSNHPITFHEGSGTFTVTFDVSYKATLSDRMLMRASSENKASSSKATP 911
901 QLELPVKYAVVAVVTSHGVSTKYNF-TASENTSRVMQHOYQVSNLQORSIPISLVFLVP 959
912 QLELPVKYAVVAVTMSRQESKTYFNFATSDKKKKEAEHYVNNLSQRLAISINFWVP 971
960 VELNQTIVWDRPQVTFPSNLSSTCHTKERLPSHSPFLAELRKAPVNVNSIAVCORIQDI 1019
972 VLLNGVAVVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRFMLDCSIADCLQPCDV 1029
1020 PFPGIOERFNATLKGNLSPDWYIKTSHNHLLIVSTAELFNDSTVTLTPGQGFVRSQTE 1079
1030 PSFSVQEEHLDFTLKGNLSPGVVRETLQKKVLVWSVAEITFDTSVYSQLPGQAFNQAQKE 1089
1080 TKVPEPPEVFNPLIVGSSVGGLLALLALITAAALYKLGFTFKRYKMMSE 1128
1090 MVLBEDEVVNAIPIIMGSSVCGALLLALITATLYKLGFTFKRYKEMLE 1138

Search completed: June 7, 2004, 17:38:55
Job time : 42.8905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds
(without alignments)
3199.127 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKROYKDMSEGGPGABPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.psp:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.psp:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.psp:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.psp:*
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.psp:*
6: /cgn2_6/ptodata/2/1aa/PTCFILES1.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5868	99.9	1153	1	US-08-173-497-3
2	5868	99.9	1153	1	US-08-286-889-3
3	5868	99.9	1153	1	US-08-485-618-3
4	5868	99.9	1153	1	US-08-362-652-3
5	5868	99.9	1153	2	US-08-605-672-3
6	5868	99.9	1153	2	US-08-482-293A-3
7	5868	99.9	1153	2	US-08-943-363-3
8	5868	99.9	1153	3	US-09-193-043-3
9	5868	99.9	1153	4	US-09-688-307A-3
10	5868	99.9	1153	4	US-09-350-259-3
11	5837.5	99.4	1152	2	US-08-476-062A-43
12	5837.5	99.4	1152	5	PCT-US96-01314-43
13	5837.5	99.4	1152	6	5424399-2
14	3473	59.1	1163	2	US-08-476-062A-44
15	3473	59.1	1163	5	PCT-US96-01314-44
16	3450	58.7	1163	1	US-08-173-497-4
17	3450	58.7	1163	1	US-08-286-889-4
18	3450	58.7	1163	1	US-08-485-618-4
19	3450	58.7	1163	1	US-08-362-652-4
20	3450	58.7	1163	2	US-08-605-672-4
21	3450	58.7	1163	2	US-08-482-293A-4
22	3450	58.7	1163	2	US-08-943-363-4
23	3450	58.7	1163	3	US-09-193-043-4
24	3450	58.7	1163	4	US-09-688-307A-4
25	3450	58.7	1163	4	US-09-350-259-4
26	3417	58.2	1161	1	US-08-173-497-2
27	3417	58.2	1161	1	US-08-286-889-2

28	3417	58.2	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3417	58.2	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3417	58.2	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3417	58.2	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3417	58.2	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3417	58.2	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3417	58.2	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3417	58.2	1161	4	US-09-350-259-2	Sequence 2, Appli
36	3401.5	57.9	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3401.5	57.9	1161	1	US-08-605-672-99	Sequence 99, Appl
38	3401.5	57.9	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3401.5	57.9	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3401.5	57.9	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3401.5	57.9	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3401.5	57.9	1161	4	US-09-350-259-99	Sequence 99, Appl
43	3239.5	55.1	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3239.5	55.1	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3239.5	55.1	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 1; Mismatches 0; Gaps 0;

OY 1 FNLDTENAMTFOENARGFGQSVVQGGSRVVVGAPOEIVAAANORGLYCCDYSTGSCPEI 60

DB 17 FNLDTENAMTFOENARGFGQSVVQGGSRVVVGAPOEIVAAANORGLYCCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGLCFLPGSNLRQPOK 120
DB 77 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGLCFLPGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXKTLFSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXKTLFSLMOYSEEP 196
QY 181 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGIRKVRRELLNITNGARKNAFKLILVI 240
DB 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGIRKVRRELLNITNGARKNAFKLILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
QY 301 NFEALKTIONQLREKI FAIEGTOTGSSSEHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKI FAIEGTOTGSSSEHEMSQEGFSAATNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVMFR 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVILGAPRYOHIGLVMFR 436
QY 421 QNTGWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLGAPHYETOTRGGOVSCPL 480
DB 437 QNTGWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLGAPHYETOTRGGOVSCPL 496
QY 481 PRGRARWQCDVLYGCGQGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNEGAVYLF 540
DB 497 PRGRARWQCDVLYGCGQGWGFAGALTVLGDVNGDKLTDVAIGAPGEEDNEGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGQDITWDGLVDLTVGAQGHYLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGQDITWDGLVDLTVGAQGHYLLRSQ 616
QY 601 PVLAVKAMEFNPREVARNVECDQVVKGEAGEVCLHVQKSTDRRLREGQIQSVVT 660
DB 617 PVLAVKAMEFNPREVARNVECDQVVKGEAGEVCLHVQKSTDRRLREGQIQSVVT 676
QY 661 YDLALDSGRPSHRAVFNKSTRTQVGLGLTQTCBTLKQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPSHRAVFNKSTRTQVGLGLTQTCBTLKQLPNCIEDPVSPIVLRNF 736
QY 721 SLVCTPLSAFGLNRPVLAEDAQRLFTALFPPEKNCNDNICQDLSITFSPMSLDCLVVG 780
DB 737 SLVCTPLSAFGLNRPVLAEDAQRLFTALFPPEKNCNDNICQDLSITFSPMSLDCLVVG 796
QY 781 GPRFNVTVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASTEV 840
DB 797 GPRFNVTVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTNIPFDVDSKASLGKLLKANVTSENNKPTNKTET 900
DB 857 SGALKSTSCSINHPIPPENSEVTNIPFDVDSKASLGKLLKANVTSENNKPTNKTET 916
QY 901 QLELPVKYAVVTVTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
DB 917 QLELPVKYAVVTVTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTPSENLSTCTHKLPSHDFLAEIRKAPVNVCSIAVCQRIQCDIP 1020
DB 977 RLNQTIVWDRPQVTPSENLSTCTHKLPSHDFLAEIRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 PFGIOEEFNATLKGNE.SFDWIKTSHNHLITVSTAEILFNDSVFTLLPGQAPVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNE.SFDWIKTSHNHLITVSTAEILFNDSVFTLLPGQAPVRSOTET 1096
QY 1081 KYPEPEVFNPLPIIVGSSVGGILLALITAALYKLGFFKRYQKDMWSEGGPPGAEPQ 1137
DB 1097 KYPEPEVFNPLPIIVGSSVGGILLALITAALYKLGFFKRYQKDMWSEGGPPGAEPQ 1153

RESULT 2
US-08-286-889-3
Sequence 3, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murtay & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: P38,659
REFERENCE/DOCKET NUMBER: 27866/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFORNAERGQGVVQLOGSRVVGAPQEIIVAANQORGLYQCDYSTGCEPI 60
DB 17 FNLDTENAMTFORNAERGQGVVQLOGSRVVGAPQEIIVAANQORGLYQCDYSTGCEPI 76
QY 61 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGLCFLPGSNLRQPOK 120
DB 77 RLQVPEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVKGLCFLPGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXKTLFSLMOYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKXKTLFSLMOYSEEP 196
QY 181 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGIRKVRRELLNITNGARKNAFKLILVI 240
DB 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGIRKVRRELLNITNGARKNAFKLILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
QY 301 NFEALKTIONQLREKI FAIEGTOTGSSSEHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKI FAIEGTOTGSSSEHEMSQEGFSAATNSGPLLSTVGSYDWAG 376

361	QY	GVPLYTSKEKSTFINMTVRVDS	DNNDAYLVAAAIILNRNRVQSLVLCAR	YOHIGLVAMFR	420
362	DB				421
363	DB				422
364	DB				423
365	DB				424
366	DB				425
367	QY	QNTGWESNANVKGTOIGAYFGASLCS	VVDVDSNGSTDVLVLI	GAPHYYQTRGGQVSVCP	480
368	DB				481
369	DB				482
370	DB				483
371	DB				484
372	DB				485
373	QY	PRGQARWQCDAVLYGEOGQWGRSFGAAL	TVLGDVNGDKLTDVAL	GAPBEDNRGAVYLP	540
374	DB				541
375	DB				542
376	DB				543
377	DB				544
378	DB				545
379	QY	HQTSOGISPSHSQRIAGSKLSPLO	QVFGOSLGGODLTMGDL	VTVAQGHVLLRSQ	6000
380	DB				6001
381	DB				6002
382	DB				6003
383	DB				6004
384	QY	PVLRVKAIEMBNPREVARNVPECNDQVVK	GEAGEVVRCLHVOKSTRDLREGOI	QSVWT	660
385	DB				661
386	DB				662
387	DB				663
388	DB				664
389	QY	YDLALDSCRPSRAVFNETKNSRRQTV	UGLGTQTCETLKL	PNCIEDPVPVILRLNF	720
390	DB				721
391	DB				722
392	DB				723
393	QY	SLVGTPLSAFGLNRPVLAEDAQR	LFTALPFPEKNCNDNI	CDDLSITFSFMSDCLLVWG	780
394	DB				781
395	DB				782
396	QY	GRFENVTVTVRNDGDSYKTQVTFPP	PLDLSYRKVSTLQNRORSRWLLACBASSTEV	840	
397	DB				841
398	DB				842
399	QY	-841	SGALKSTSCSINHPIFPENSEVFNITFDV	DSKASLGNKLLKANVTSENNMPRTNTEF	900
400	DB				901
401	QY	QLELPVKYAVTWVTSRGVTKYLNF	TASENTRVWQHQYQVSNLQBSLPI	SISLVFLVPV	960
402	DB				961
403	DB				962
404	QY	QLELPVKYAVTWVTSRGVSTKYLNF	TASENTRVWQHQYQVSNLQBSLPI	SISLVFLVPV	976
405	DB				977
406	QY	RLNQTVIWDPRQVTFSENLSSTCHT	KERLPSSHDFLAELRKAPVWNC	SIAYCQIQCDIP	1020
407	DB				1021
408	DB				1022
409	QY	PFQIOEEFNATLKNGLSPDWYIKTS	NNHLLIVSTABILENDSVFTLL	PGQAGFVRSOTET	1080
410	DB				1081
411	DB				1082
412	QY	KVEPPEVNPVPLIVGSSVGLLILLAL	ITAALYKLGFFKQTKNMWSSGGP	PGPAEPQ	1137
413	DB				1138
414	DB	KVEPPEVNPVPLIVGSSVGLLILLAL	ITAALYKLGFFKQTKNMWSSGGP	PGPAEPQ	1153

257	Db	TDGEKFGDPLGYEDVLP	ADREGVIRYVIGVDA	FRSEKSRQELNT	IASKPPDRHVFQN	316
301	Qy	NFEALKTQNLQREKIP	ATEGTQTCSSSFHEMS	QEGFSAITSNGPL	STVGSVDNAG	360
317	Db	NFEALKTQNLQREKIP	ATEGTQTCSSSFHEMS	QEGFSAITSNGPL	STVGSVDNAG	376
361	Qy	GVPLYTSKEKSTFIN	NTRVDSMDNDAYLGY	AAAAIILRNRVQSL	VLGAPRYQHIGLVAMFR	420
377	Db	GVPLYTSKEKSTFIN	NTRVDSMDNDAYLGY	AAAAIILRNRVQSL	VLGAPRYQHIGLVAMFR	436
421	Qy	QNTGHWESNANVKG	TQIGAYFGALSCSDVD	SDNGSTDLVLIGAPHY	TECTRGQGVSVCP	480
437	Db	QNTGHWESNANVKG	TQIGAYFGALSCSDVD	SDNGSTDLVLIGAPHY	TECTRGQGVSVCP	496
481	Qy	PRGQARQWCDAVL	YGEQGQPMGRFGAAL	TVLGDVNGDKLT	DVAI.GAPCEEDNRGA	540
497	Db	PRGQARQWCDAVL	YGEQGQPMGRFGAAL	TVLGDVNGDKLT	DVAI.GAPCEEDNRGA	556
541	Qy	HGTSGSGISPSHS	QRISQAKSLSPRIQY	FGQSLSGSGDLT	MDGLVDLTVGAOGFVLL	600

```

1  RESULT 3
2  US-08-485-618-3
3  US-08-485-618-3 Application US/08485618
4  Sequence 3, Application US/08485618
5  Patent No. 5728533
6  Patent No. 5728533
7  GENERAL INFORMATION:
8  APPLICANT: Gallatin, W. Michael
9  APPLICANT: Van der Vieren, Monica
10 TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
11 TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
12 NUMBER OF SEQUENCES: 103
13 NUMBER OF SEQUENCES: 103
14 CORRESPONDENCE ADDRESS:
15 ADDRESSSE: Marshall, O'Toole, Gerstein, Murray & Borun
16 STREET: 233 South Wacker Drive, 6300 Sear Tower
17 CITY: Chicago
18 STATE: Illinois
19 COUNTRY: United States
20 ZIP: 60608-6402
21 ZIP: 60608-6402
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Floppy disk
24 COMPUTER: IBM PC compatible
25 COMPUTER: IBM PC compatible

```

557 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVCAQGHVLLLRQ 616
601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 660
617 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 676
661 YDLALDSGRPHSRAVNETKNSRRTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
677 YDLALDSGRPHSRAVNETKNSRRTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
721 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPFPEKXKCGNDNICQDDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPFPEKXKCGNDNICQDDLSITFSFMSLDCLVVG 796
781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACESASSTEV 840
797 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACESASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKLLKANTYSENMPRTNKEP 900
857 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKLLKANTYSENMPRTNKEP 916
901 QLELPVKIAYVMVTSKGVSTKYNFTASNTSRVMOHQVOVNLGORSIPISLVFLVPV 960
917 QLELPVKIAYVMVTSKGVSTKYNFTASNTSRVMOHQVOVNLGORSIPISLVFLVPV 976
961 RLNQTVIWRDPQVTFPSNLSSTCHTKERLPSHSDFLAELKAPVNCSTIAVCORIQCIP 1020
977 RLNQTVIWRDPQVTFPSNLSSTCHTKERLPSHSDFLAELKAPVNCSTIAVCORIQCIP 1036
1021 FFGIOEPNATLKNLSPPDKVTKTSHNHLIVSTABILENDSVFTLLPGOGAFVRSOTET 1080
1037 FFGIOEPNATLKNLSPPDKVTKTSHNHLIVSTABILENDSVFTLLPGOGAFVRSOTET 1096
1081 KVEPFFVNPPLPLTVGSSVGGILLALLITAAALYKLGFFKQYKDMMSSEGGPPCAEPQ 1137
1097 KVEPFFVNPPLPLTVGSSVGGILLALLITAAALYKLGFFKQYKDMMSSEGGPPCAEPQ 1153

RESULT 4
US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.9%; Score 5868; DB 1; length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTPOENARGFGQSVVQLQSSRVVVGAPQELVAANQORSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTPOENARGFGQSVVQLQSSRVVVGAPQELVAANQORSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHTCSENTYVKGCLFLFGNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHTCSENTYVKGCLFLFGNLRQPOK 136
QY 121 PPEARLGGCPOEDSDIAPLIDGSGSII PHDFRMKEFVSTVMEQLKSKTLPFLMOYSEBF 180
DB 137 PPEARLGGCPOEDSDIAPLIDGSGSII PHDFRMKEFVSTVMEQLKSKTLPFLMOYSEBF 196
QY 181 RIHFTFKGQNNPNRSLVKPITOLLGRTHATATGVRKIRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKGQNNPNRSLVKPITOLLGRTHATATGVRKIRELLNITNGARKNAFKILIVI 256
QY 241 TDGKFGDPLGHEDVTPADREGVIRYVIGVGDAPFRSEKSRQELNITASKPRDHVFN 300
DB 257 TDGKFGDPLGHEDVTPADREGVIRYVIGVGDAPFRSEKSRQELNITASKPRDHVFN 316
QY 301 NFEALKTIONQLREKIPAIETGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIPAIETGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVPLTYSKEKSTFINMTRVDSMDNDAYLGYAAAIIRNRVQSLVLCAPYQHIGLVAMER 420
DB 377 GVPLTYSKEKSTFINMTRVDSMDNDAYLGYAAAIIRNRVQSLVLCAPYQHIGLVAMER 436
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQVSVCP 480
DB 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQVSVCP 496
QY 481 PRGQARWQCDVLYGEGQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEENRGAVYLF 540
DB 497 PRGQARWQCDVLYGEGQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVCAQGHVLLLRQ 600
DB 557 HGTSGSGISPHSQRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVCAQGHVLLLRQ 616
QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 660
DB 617 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRAVNETKNSRRTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRAVNETKNSRRTQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPFPEKXKCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEDAQRLLFTALPFPEKXKCGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1036
Qy 1021 PFGIOSEFNATLGNLSFQWYKTSNNHLLIYSTABILFNDVSFTLLPCGGAPVRSOTET 1080
Db 1037 PFGIOSEFNATLGNLSFQWYKTSNNHLLIYSTABILFNDVSFTLLPCGGAPVRSOTET 1096
Qy 1081 KVEPFVFPNPLPIVGVSSVGGLLLLALITAAALYKLGFFRKQYKDMSEGPPGABEPQ 1137
Db 1097 KVEPFVFPNPLPIVGVSSVGGLLLLALITAAALYKLGFFRKQYKDMSEGPPGABEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTQENARPGQSVQVQLQGSRVVVGAPQEIIVAAQNRGSLVOCYDSTGSCBPI 60
Db 17 FNLDTENAMTQENARPGQSVQVQLQGSRVVVGAPQEIIVAAQNRGSLVOCYDSTGSCBPI 76
Qy 61 RLOVPVEAVNWSLGLSLAATTPOLLACGPTVHCTCSENTVVKGLCLFGLFNSNLQOQPOK 120
Db 77 RLOVPVEAVNWSLGLSLAATTPOLLACGPTVHCTCSENTVVKGLCLFGLFNSNLQOQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEFVSTVMEQLKSKTLPFSLMOYSBEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFFRMKEFVSTVMEQLKSKTLPFSLMOYSBEF 196
Qy 181 RIHPTFKFQONPNRSLVKITQLLGRTHATATGVRKVIKRELLNITNGARKNAKILVI 240
Db 197 RIHPTFKFQONPNRSLVKITQLLGRTHATATGVRKVIKRELLNITNGARKNAKILVI 256
Qy 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQV 316
Qy 301 NFEALKTIQNLREKIFAIETGTQSSSSFEHEMSQEGFSAAITNGPILLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIETGTQSSSSFEHEMSQEGFSAAITNGPILLSVGSYDWAG 376
Qy 361 GYPLVTSKEKSTFINNTRVDSMDNDAYLGAAAILRNKRVSLVGLGAPRYQHIGLVAMFR 420
Db 377 GYPLVTSKEKSTFINNTRVDSMDNDAYLGAAAILRNKRVSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQSVSCPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQSVSCPL 496
Qy 481 PRGQARWQCDAYLYGEOGQPKRGGAALTVLGDVNGDKLTDVAIGAPEEDNRGAVLIF 540
Db 497 PRGQARWQCDAYLYGEOGQPKRGGAALTVLGDVNGDKLTDVAIGAPEEDNRGAVLIF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSSLSGGQDLTMDGLVLTGVAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSSLSGGQDLTMDGLVLTGVAQGHVLLLRQ 616
Qy 601 PVLRVKAIMBEPNPREVARNVFECDQVYKGEAGVRVCLHVOKSTRDLRGQIQSVVT 660
Db 617 PVLRVKAIMBEPNPREVARNVFECDQVYKGEAGVRVCLHVOKSTRDLRGQIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNENKSTRTOVLGLTQTCETLKLQLPNCIBDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNENKSTRTOVLGLTQTCETLKLQLPNCIBDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLFTALPPFKNGCNDNICQDDLSITPSFMSLCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAORLFTALPPFKNGCNDNICQDDLSITPSFMSLCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORSQBSWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1036

QY 1021 PFGIOEFNATLKNLSFDWYIKTSNHNLLIVSTABILFNDVSFTLLPQCGAFVRSOTET 1080
DB 1037 PFGIOEFNATLKNLSFDWYIKTSNHNLLIVSTABILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 KVEPFVEPNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1097 KVEPFVEPNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1153

RESULT 6
US-08-482-293A-3
: Sequence 3, Application US/08482293A
: Patent No. 5831029
: GENERAL INFORMATION:
: APPLICANT: Gallatin, W. Michael
: APPLICANT: Van der Vieren, Monica
: TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
: NUMBER OF SEQUENCES: 103
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive, 6300 Sear Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION NUMBER: US/08/482,293A
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/173,497
: FILING DATE: 23-DEC-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/286,889
: FILING DATE: 5-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/362,652
: FILING DATE: 21-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: 36,659
: REFERENCE/DOCKET NUMBER: 27866/32684
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1153 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDENAMTQENARGPQGSVWVVGAPQIIVANQSGSYQCDYSTGSCPEI 60
DB 17 FNLDENAMTQENARGPQGSVWVVGAPQIIVANQSGSYQCDYSTGSCPEI 76

QY 61 RLQPVVEAVNMSGLSLAATTSPTQLACGPTVHTQCSNTYVKGCLFLGCSNLRQOPQK 120
DB 77 RLQPVVEAVNMSGLSLAATTSPTQLACGPTVHTQCSNTYVKGCLFLGCSNLRQOPQK 136

QY 121 PFALRGCTQEDSDIAFLIDSGSIIIPDFRMKEFVSTWELKSKTLPFLMQYSEEP 180

DB 137 PFALRGCTQEDSDIAFLIDSGSIIIPDFRMKEFVSTWELKSKTLPFLMQYSEEP 196
QY 181 RIHFTFKPQNNPNRSLVKPITOLLGRTHATGVRKIVIRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKPQNNPNRSLVKPITOLLGRTHATGVRKIVIRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDRHVQFN 300
DB 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDRHVQFN 316
QY 301 NPEALKTIQNLREKIPAIETGTQSSSSFEHMSQEGFSAAITNGHLLSTVGSYDNAG 360
DB 317 NPEALKTIQNLREKIPAIETGTQSSSSFEHMSQEGFSAAITNGHLLSTVGSYDNAG 376
QY 361 GYPLVTSKEKSTFINMTRVDSMDNDAYLGAYAAIILNRVQSLVLGARYQHIGLVAMFR 420
DB 377 GYPLVTSKEKSTFINMTRVDSMDNDAYLGAYAAIILNRVQSLVLGARYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASICSDVDVDSNGSTDVLVIGAPHYETRGQGVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASICSDVDVDSNGSTDVLVIGAPHYETRGQGVSVCLP 496
QY 481 PRGQARWQCDAYLYGEOGQPMWRFGAALTVDVNGDKLTDAVIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDAYLYGEOGQPMWRFGAALTVDVNGDKLTDAVIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTWGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTWGAQGHVLLRSQ 616
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRGQIQSVVT 660
DB 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLRGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLRPVLADAOBLPTALPPPEKNCNDNICODDLSITFSMSLDCLVVG 780
DB 737 SLVGTPLSAFGLRPVLADAOBLPTALPPPEKNCNDNICODDLSITFSMSLDCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPPFDLSYRKVSTLQNRQSRMSRLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSVRTQVTFPPFDLSYRKVSTLQNRQSRMSRLACESASSTEV 856
QY 841 SGALKSTCSINHPIPPENSEVTFTDVSQKSLGKLLKANTSENNPRTNKTFE 900
DB 857 SGALKSTCSINHPIPPENSEVTFTDVSQKSLGKLLKANTSENNPRTNKTFE 916
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIOCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIOCDIP 1036
QY 1021 PFGIOEFNATLKNLSFDWYIKTSNHNLLIVSTABILFNDVSFTLLPQCGAFVRSOTET 1080
DB 1037 PFGIOEFNATLKNLSFDWYIKTSNHNLLIVSTABILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 KVEPFVEPNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1097 KVEPFVEPNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1153

RESULT 7
US-08-943-363-3
: Sequence 3, Application US/08943363
: Patent No. 5837478
: GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
 APPLICANT: Van der Vieren, Monica
 TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
 NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 233 South Wacker Drive, 6300 Sear Tower
 City: Chicago

STATE: Illinois
 COUNTRY: United States
 ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,363

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,652

FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32684

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI	76
QY	61	RLOQVPEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK	120
DB	77	RLOQVPEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK	136
QY	121	FPEALRGCPQEDSDIAFLDGSIIIPDPRMKEFVSTVMEQLKSKTLFSLMQSYSEEF	180
DB	137	FPEALRGCPQEDSDIAFLDGSIIIPDPRMKEFVSTVMEQLKSKTLFSLMQSYSEEF	196
QY	181	RIHFTPEFQNNPRSLVKPIPOLGRHTATGVRKVIRELNIITNGARKNAFKILWVI	240
DB	197	RIHFTPEFQNNPRSLVKPIPOLGRHTATGVRKVIRELNIITNGARKNAFKILWVI	256
QY	241	TDGEKFGDPLGYEDVDEADREGVIRVVGDAFRSEKSRQBLNTIASKPPRDHVFQVN	300
DB	257	TDGEKFGDPLGYEDVDEADREGVIRVVGDAFRSEKSRQBLNTIASKPPRDHVFQVN	316
QY	301	NFEALKTIQNLREKIFAISGTQYTGSSSSPEHEMSQEGPSAAITSGNPLISTVGSYDWAG	360
DB	317	NFEALKTIQNLREKIFAISGTQYTGSSSSPEHEMSQEGPSAAITSGNPLISTVGSYDWAG	376

QY	361	GVFLYTSKSKSTFINMTRVDSMDMDAYLGAAAILLRNVQSLVGLGAPRYQHIGLIVAMFR	420
DB	377	GVFLYTSKSKSTFINMTRVDSMDMDAYLGAAAILLRNVQSLVGLGAPRYQHIGLIVAMFR	436
QY	421	QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSSTDLVILGAPHYHBTQFGGVSVCPL	480
DB	437	QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSSTDLVILGAPHYHBTQFGGVSVCPL	496
QY	481	PRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGEDNKGAVLYF	540
DB	497	PRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNVDGKLTVDVAIGAPGEDNKGAVLYF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAOGHVLILLASQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAOGHVLILLASQ	616
QY	601	PVLRYKAIEMENPREVARNVFECDNDVVVKGKAGEVRVCLHVQKSTRDLREGGIIQS VVT	660
DB	617	PVLRYKAIEMENPREVARNVFECDNDVVVKGKAGEVRVCLHVQKSTRDLREGGIIQS VVT	676
QY	661	YDLALDSGRPHSRVAFNETKNSTRRTQVGLGTQTCETLKLQLPNCIEDDPVSPVLRNLF	720
DB	677	YDLALDSGRPHSRVAFNETKNSTRRTQVGLGTQTCETLKLQLPNCIEDDPVSPVLRNLF	736
QY	721	SLVGTPLSAGPNLRPVLAEADAQRLFTALFPPEKKNCDNII CODDLSITFSFMSLDCLVVG	780
DB	737	SLVGTPLSAGPNLRPVLAEADAQRLFTALFPPEKKNCDNII CODDLSITFSFMSLDCLVVG	796
QY	781	GPREFNVTTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSRLACESASSTEV	840
DB	797	GPREFNVTTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSRLACESASSTEV	856
QY	841	SGALKSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF	900
DB	857	SGALKSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF	916
QY	901	QLELPVKYAVVYVTVSHGVSTKYLNFTASENTRVMQHOYOVSNLQSRSLPISLVFLVPV	960
DB	917	QLELPVKYAVVYVTVSHGVSTKYLNFTASENTRVMQHOYOVSNLQSRSLPISLVFLVPV	976
QY	961	RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLABELKAPVYVNCISLAVCORIQCDIP	1020
DB	977	RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLABELKAPVYVNCISLAVCORIQCDIP	1036
QY	1021	FFGIOEEFNATLKGSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQOGAFVRSQDET	1080
DB	1037	FFGIOEEFNATLKGSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQOGAFVRSQDET	1096
QY	1081	KVEPEVFNPLPLIVGSSVGGILLALITALYKLGFFKQYKQVKSXSGGPPGABPQ	1137
DB	1097	KVEPEVFNPLPLIVGSSVGGILLALITALYKLGFFKQYKQVKSXSGGPPGABPQ	1153

RESULT 8

US-09-193-043-3

Sequence 3, Application US/09193043

Patent No. 6251395

GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.

APPLICANT: Van der Vieren, Monica

TITLE OF INVENTION: No. 6251395e1 Human 2

FILE REFERENCE: 27866/35004

CURRENT APPLICATION NUMBER: US/09/193,043

CURRENT FILING DATE: 1998-11-16

EARLIER APPLICATION NUMBER: 08/173,497

EARLIER FILING DATE: 1993-12-23

EARLIER APPLICATION NUMBER: 08/286,889

EARLIER FILING DATE: 1994-08-05

EARLIER APPLICATION NUMBER: 08/362,652

EARLIER FILING DATE: 1994-12-21

EARLIER APPLICATION NUMBER: 08/943,363

EARLIER FILING DATE: 1997-10-03

```

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match          99.9%; Score 5868; DB 3; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQGSLSYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQGSLSYQCDYSTGSCBPI 76
QY 61 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFGLGSLNRQOQOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFGLGSLNRQOQOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQLKKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQLKKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNPRSLVKPITQLLGRTHATGVRKVIKRLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNPRSLVKPITQLLGRTHATGVRKVIKRLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGAFSEKSRQELNTIASKPRDHVQFN 300
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGAFSEKSRQELNTIASKPRDHVQFN 316
QY 301 NFEALKTQNLREKIPAEIGTQTQSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 360
DB 317 NFEALKTQNLREKIPAEIGTQTQSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 376
QY 361 GYFLYTSKSKSPINMTRVDSMDNDAYLGAAAIILNRVQSVILGAPYCHIGLVAMFR 420
DB 377 GYFLYTSKSKSPINMTRVDSMDNDAYLGAAAIILNRVQSVILGAPYCHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQSVCP 496
QY 481 PRQORARWQCDAYLGEQOPWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDAYLGEQOPWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 616
QY 601 PVLVRVKAIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQOQSVVT 660
DB 617 PVLVRVKAIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQOQSVVT 676
QY 661 YDLALDSGRPHSAVNETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVNETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736
QY 721 SLVGTPLSAFGRNLPVLAEDAQLFTALPFPEKNCNDNIQQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGRNLPVLAEDAQLFTALPFPEKNCNDNIQQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVTRNDGDSYRTQTFPEPLDLSYRKVSTLQNSQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTRNDGDSYRTQTFPEPLDLSYRKVSTLQNSQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTF 900
DB 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTF 916
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSLSISLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSLSISLVFLVPV 976
QY 961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCORLOCDIP 1020
DB 977 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCORLOCDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1096
QY 1081 KVBEPFEVNPPLHIVGSSVGGILLALITAAALYKLGFKFKQYKDMSEGGPPGAEPO 1137
DB 1097 KVBEPFEVNPPLHIVGSSVGGILLALITAAALYKLGFKFKQYKDMSEGGPPGAEPO 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36846
; CURRENT APPLICATION NUMBER: US/09/688, 307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193, 043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605, 672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173, 497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286, 889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362, 652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943, 363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match          99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQGSLSYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVAAANQGSLSYQCDYSTGSCBPI 76
QY 61 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFGLGSLNRQOQOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHTQCSNTYVVKGLCFGLGSLNRQOQOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQLKKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKKEFVSTVMEQLKKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNPRSLVKPITQLLGRTHATGVRKVIKRLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNPRSLVKPITQLLGRTHATGVRKVIKRLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGAFSEKSRQELNTIASKPRDHVQFN 300
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGAFSEKSRQELNTIASKPRDHVQFN 316
QY 301 NFEALKTQNLREKIPAEIGTQTQSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 360
DB 317 NFEALKTQNLREKIPAEIGTQTQSSSFHEMSQEGFSAAITSNGLPILSTVGSYDNAG 376
QY 361 GYFLYTSKSKSPINMTRVDSMDNDAYLGAAAIILNRVQSVILGAPYCHIGLVAMFR 420
DB 377 GYFLYTSKSKSPINMTRVDSMDNDAYLGAAAIILNRVQSVILGAPYCHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQSVCP 496
QY 481 PRQORARWQCDAYLGEQOPWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRQORARWQCDAYLGEQOPWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGQAQGHVILLRSQ 616
QY 601 PVLVRVKAIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQOQSVVT 660
DB 617 PVLVRVKAIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRQOQSVVT 676
QY 661 YDLALDSGRPHSAVNETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVNETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736
QY 721 SLVGTPLSAFGRNLPVLAEDAQLFTALPFPEKNCNDNIQQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGRNLPVLAEDAQLFTALPFPEKNCNDNIQQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVTRNDGDSYRTQTFPEPLDLSYRKVSTLQNSQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTRNDGDSYRTQTFPEPLDLSYRKVSTLQNSQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTF 900
DB 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTF 916
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSLSISLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHQVQVSNLQORSLSISLVFLVPV 976
QY 961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCORLOCDIP 1020
DB 977 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCORLOCDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPFGQAFVRSOTET 1096
QY 1081 KVBEPFEVNPPLHIVGSSVGGILLALITAAALYKLGFKFKQYKDMSEGGPPGAEPO 1137
DB 1097 KVBEPFEVNPPLHIVGSSVGGILLALITAAALYKLGFKFKQYKDMSEGGPPGAEPO 1153
```

Db 317 NFEALKTQNLREKIFALEGTQTSSTSSFEHEMSOEGFSAITNGPILLSTVSGYDWMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOIHGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOIHGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 496
Qy 481 PRQORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIWEPNPREVARNVFECNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
Db 617 PVLRVKAIWEPNPREVARNVFECNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSAFVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPSIVLRNPF 720
Db 677 YDLALDSGRPHSAFVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPSIVLRNPF 736
Qy 721 SLVGTPLSAPGNLRPVLAEDAORLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLRPVLAEDAORLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWRLACSSASTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWRLACSSASTEV 856
Qy 841 SGALKSTSCSINHPPIPFENSEVTFNFTFDVDSKASLGNKLLKXANTSENMMPTNKTEF 900
Db 857 SGALKSTSCSINHPPIPFENSEVTFNFTFDVDSKASLGNKLLKXANTSENMMPTNKTEF 916
Qy 901 QLELPVKYAVYVTVSHGVSTKYLNTFASNTSRVMOHQVQVNLGORSIPISLVLVPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNTFASNTSRVMOHQVQVNLGORSIPISLVLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIQDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIQDIP 1036
Qy 1021 PFCIOEPENATLKNLSFDWYKTSNNHLIVSTAELLFNDVSFTLLPGOGAFVRSQTE 1080
Db 1037 PFCIOEPENATLKNLSFDWYKTSNNHLIVSTAELLFNDVSFTLLPGOGAFVRSQTE 1096
Qy 1081 KVPFPFEPNPLPIVSGSVGLLALLITALYKLGFFKQYKXDMWSEGGPPGABPQ 1137
Db 1097 KVPFPFEPNPLPIVSGSVGLLALLITALYKLGFFKQYKXDMWSEGGPPGABPQ 1153

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350.259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193.043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173.497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286.889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362.652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943.363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PST
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQVQGGSRVVVGAPQBEIVAAANQORSLVQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQVQGGSRVVVGAPQBEIVAAANQORSLVQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCLFLGSLNLRQOPQK 120
Db 77 RLQVPEAVNNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCLFLGSLNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAPLIDSGSIIPHDPRMKFVSTVMEOLKSKTLPISLMQYSBEF 180
Db 137 FPEALRGCPQEDSDIAPLIDSGSIIPHDPRMKFVSTVMEOLKSKTLPISLMQYSBEF 196
Qy 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVI RELLNITNGARKNAFKILYVI 240
Db 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVI RELLNITNGARKNAFKILYVI 256
Qy 241 TDGEKPGDPLGYEDVIPEADREGVIRYIVGVGDAFRSEKSRQELNLTASKPRDHVQVN 300
Db 257 TDGEKPGDPLGYEDVIPEADREGVIRYIVGVGDAFRSEKSRQELNLTASKPRDHVQVN 316
Qy 301 NFEALKTQNLREKIFALEGTQTSSTSSFEHEMSOEGFSAITNGPILLSTVSGYDWMAG 360
Db 317 NFEALKTQNLREKIFALEGTQTSSTSSFEHEMSOEGFSAITNGPILLSTVSGYDWMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOIHGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYOIHGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 496
Qy 481 PRQORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQORARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIWEPNPREVARNVFECNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
Db 617 PVLRVKAIWEPNPREVARNVFECNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSAFVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPSIVLRNPF 720
Db 677 YDLALDSGRPHSAFVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPSIVLRNPF 736
Qy 721 SLVGTPLSAPGNLRPVLAEDAORLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLRPVLAEDAORLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWRLACSSASTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRWRLACSSASTEV 856
Qy 841 SGALKSTSCSINHPPIPFENSEVTFNFTFDVDSKASLGNKLLKXANTSENMMPTNKTEF 900

857 SGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRINTKTEP 916
 901 QLELPVKYAVYMYVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 960
 917 QLELPVKYAVYMYVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 976
 961 RLNQTVWDRPQVTSSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVCORIQCDIP 1020
 977 RLNQTVWDRPQVTSSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVCORIQCDIP 1036
 1021 FFGIOEENATLKGMLSFDMYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1080
 1037 FFGIOEENATLKGMLSFDMYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1096
 1081 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKROYKMGSEGGPPGAE 1137
 1097 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKROYKMGSEGGPPGAE 1153

RESULT 1:

US-08-476-062A-43
 : Sequence 43, Application US/08476062A
 : Patent No. 5877275
 : GENERAL INFORMATION:

: APPLICANT: Attnout, M. Amin
 : TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 : NUMBER OF SEQUENCES: 53
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Fish & Richardson P.C.
 : STREET: 225 Franklin Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: US
 : ZIP: 02110-2804

COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: Windows95
 : SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/476,062A
 : FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 08/216,081
 : FILING DATE: 21-MAR-1994
 : APPLICATION NUMBER: 07/537,830
 : FILING DATE: 04-JAN-1991
 : APPLICATION NUMBER: 07/539,842
 : FILING DATE: 18-JUN-1990
 : APPLICATION NUMBER: 07/212,573
 : FILING DATE: 28-JUN-1988

ATTORNEY/AGENT INFORMATION:

: NAME: Freeman, John W.
 : REGISTRATION NUMBER: 29,066
 : REFERENCE/DOCKET NUMBER: 00786/068003
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 617/542-5070
 : TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 43:

: SEQUENCE CHARACTERISTICS:
 : LENGTH: 1152 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : FRAGMENT TYPE: internal

US-08-476-062A-43

Query Match 99.4%; Score 5837.5; DB 2; Length 1152;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1130; Conservative 3; Mismatches 1; Gaps 1;

1 FNLDTENAMTFOENARGFGOSVVQLOGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCBPI 60
 17 FNLDTENAMTFOENARGFGOSVVQLOGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCBPI 76
 61 RLQVPVEAVNMSLGLSLAATTPPQOLLAGCPTVHTQCSNTYVKGCLCFPLGNSNRQOQOK 120
 77 RLQVPVEAVNMSLGLSLAATTPPQOLLAGCPTVHTQCSNTYVKGCLCFPLGNSNRQOQOK 136
 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKSKSLFSLMOYSSEF 180
 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKSKSLFSLMOYSSEF 196
 181 RHFTFKFQONPNRSLVKDITOLLGHTHTATGVRKVIIRLLNITNGARKNAKILIVI 240
 197 RHFTFKFQONPNRSLVKDITOLLGHTHTATGVRKVIIRLLNITNGARKNAKILIVI 256
 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKEPRDHVQVN 300
 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKEPRDHVQVN 316
 301 NEFALKTIONOLREKI PAIEGTOTGSSSSPEHMSQBSAAITNSGPLLSTVGSYDWA 360
 317 NEFALKTIONOLREKI PAIEGTOTGSSSSPEHMSQBSAAITNSGPLLSTVGSYDWA 376
 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVA 420
 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVA 436
 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQVSVCP 480
 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEOTRGQVSVCP 496
 481 PRGQARWQCDVLYGSGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLP 540
 497 PRG-RARWQCDVLYGSGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLP 555
 541 HGTSGSGLSPSHSRIAGSKLSPLQYFGOSLSGGDLTMDGLVDLTVGAGCHVLLRSQ 600
 556 HGTSGSGLSPSHSRIAGSKLSPLQYFGOSLSGGDLTMDGLVDLTVGAGCHVLLRSQ 615
 601 PVLRYKALMEENPREVARNVFECDQVVKGEAGEVRVCLFHVKQKSTRDRLEGGIQSVVT 660
 616 PVLRYKALMEENPREVARNVFECDQVVKGEAGEVRVCLFHVKQKSTRDRLEGGIQSVVT 675
 661 YDLALDSGRPHSRAVFNENKSTROTQVGLTCTCTETLKLQPLNCIEDPVSPIVRLNF 720
 676 YDLALDSGRPHSRAVFNENKSTROTQVGLTCTCTETLKLQPLNCIEDPVSPIVRLNF 735
 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
 736 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 795
 781 GPREFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQONORSORSWRLACSSASTEV 840
 796 GPREFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQONORSORSWRLACSSASTEV 855
 841 SGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRINTKTE 900
 856 SGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRINTKTE 915
 901 QLELPVKYAVYMYVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 960
 916 QLELPVKYAVYMYVSHGYSKYLNFTASENISRVMQHQYQVSNLQORSLSPLSLVPLVPV 975
 961 RLNQTVWDRPQVTSSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVCORIQCDIP 1020
 976 RLNQTVWDRPQVTSSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVCORIQCDIP 1035
 1021 FFGIOEENATLKGMLSFDMYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1080
 1036 FFGIOEENATLKGMLSFDMYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1095
 1081 KVEPFEVNPFLPIVGVSSVGGLLLLALITAAALYKLGFFKROYKMGSEGGPPGABPQ 1137

Db 1096 KVEPEVFNPLVGVSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPGABPQ 1152

RESULT 12
PCT-US96-01314-43
Sequence 43, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
TITLE OF INVENTION: ANTAGONISTS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.4%; Score 5837.5; DB 5; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGPQSVVQLQGSRRVVVGAPQEI VAAHQSGSLYQCDYSTGSCPI 60
Db 17 FNLDTENAMTFOENARGPQSVVQLQGSRRVVVGAPQEI VAAHQSGSLYQCDYSTGSCPI 76

QY 61 RLQVPVEANMSIGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSLNLCQPOK 120
Db 77 RLQVPVEANMSIGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSLNLCQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 196

QY 181 RHPTFKPEQNNPNRSLVKPTQLLGRTHATGKVRKVIKRELLNTINGARKNAFKILVI 240
Db 197 RHPTFKPEQNNPNRSLVKPTQLLGRTHATGKVRKVIKRELLNTINGARKNAFKILVI 256

QY 241 TGEKFGDPLGVEDVTPEDRGEVIRYVGVGDAPFSEKSRQELNTIASKPRDHVQVFN 300
Db 257 TGEKFGDPLGVEDVTPEDRGEVIRYVGVGDAPFSEKSRQELNTIASKPRDHVQVFN 316

QY 301 NFEALKTIONQAREKIFAIEGTQGTSSSFEHMSQEGFSAAITNGPILLSTVGVSDWAG 360
Db 317 NFEALKTIONQAREKIFAIEGTQGTSSSFEHMSQEGFSAAITNGPILLSTVGVSDWAG 376

QY 361 GVFLYTSKEKSTFNTMTVDSMDNDAYLGYAAAILNRVQSLVGLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFNTMTVDSMDNDAYLGYAAAILNRVQSLVGLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEQTRGGQSVQVCL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEQTRGGQSVQVCL 496

QY 481 PRGQARWQCDVLYGEGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGBEDNRGAVILF 540
Db 497 PRG-RARWQCDVLYGEGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGBEDNRGAVILF 555

QY 541 HGTSGSISPSHSORIASKLSPLQYFGQSLSGSDLTMDGLVLTVAQGHVLLLSQ 600
Db 556 HGTSGSISPSHSORIASKLSPLQYFGQSLSGSDLTMDGLVLTVAQGHVLLLSQ 615

QY 601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 616 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 675

QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGIOTCETLKLQNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTVLGIOTCETLKLQNCIEDPVSPIVLRNF 735

QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKXGNDNI CDDDLSTITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKXGNDNI CDDDLSTITFSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRSGRSLACESASSTEV 840
Db 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRSGRSLACESASSTEV 855

QY 841 SGALKSTSCSINHPIFPENSESVTNIITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSESVTNIITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 915

QY 901 QLELPVKYAVVMVTSVSHGVSTKYNFTASENTSRVMQHYQVSNLQSRSLPISLVLVFPV 960
Db 916 QLELPVKYAVVMVTSVSHGVSTKYNFTASENTSRVMQHYQVSNLQSRSLPISLVLVFPV 975

QY 961 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCIAVQRIQCQDIP 1020
Db 976 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVNVNCIAVQRIQCQDIP 1035

QY 1021 PFGIQEBENATLKGNLSPDWYIKTSHNHLIVSTABILFNDSPVTLPCQCAFVSQET 1080
Db 1036 PFGIQEBENATLKGNLSPDWYIKTSHNHLIVSTABILFNDSPVTLPCQCAFVSQET 1095

QY 1081 KVEPEVFNPLVGVSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPGABPQ 1137
Db 1096 KVEPEVFNPLVGVSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPGABPQ 1152

RESULT 13
5424399-2
Patent No. 5424399
APPLICANT: ARNAOUT, M. AMIN
TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/78,871
FILING DATE: 16-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 212,573
FILING DATE: 28-JUN-1988
SEQ ID NO: 2;
LENGTH: 1152
5424399-2

Query Match 99.4%; Score 5837.5; DB 6; Length 1152;

541	QY	541	HCTSGSGISPHSHORISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLIVGAGQHVVLLRSQ	600
557	DB	557	HGVLGSPS18FSPHSORISAGSKLSSRLQYFGQALSGQDLTQGLVDLIVGARGQVLLLRTR	616
601	QY	601	PVLVRKAIMBENPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTEDRLRREGIOISVWT	660
617	DB	617	PVLIVGVSQWQFIPAEIPRSAPFECRQVVSQTLVQSNICLIYDKRSKNLKGSRDLQSSVT	676
661	QY	661	YDLALDSGRPHSRAVENETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	720
677	DB	677	LDLALDPGRLSPRATFOETKNRSLSRVRVLGKAHCENFNLLPSCVEDSVPTITLINF	736
721	QY	721	SLVGTPLSAFGNLRPVLAEADAQRLFTALPFPEKXGNDNICTODLSTITPFMSLDCLVWG	780
737	DB	737	TLVGKFLLAFLNLRPLMLAAQRYFTASLPFEXKCGADHICQNLGITSFSPFGKLKSLVG	796
781	QY	781	GPREFNVTVTVRNDGDSYRTQVTFPFLDLSYKVKYSTLQNRSSQSRWRLACSSASSTEV	840
797	DB	797	SNLELNAEYVWVNDGDSYGTITTFSPAGLSYRYVAEGQOQGLSLHLTCDOSAPV--	854
841	QY	841	SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKXANTSENMPRTNKTEP	900
855	DB	855	SQGTWSTSCRIINHIIFRGGAQITFLATEDVSPRAVLGRDLLLTANVSSENNTERTSKTIF	914
901	QY	901	QLELPKYAVYVWTVTSHGVSHTKILNITAS-ENTSRVMOHQYQVSNLQORSPLPSLVLVLP	959
915	DB	915	QLELPKYAVYTVWSHEQFTKYNLNFSESEBESHVAMHRYQVNNLQORDLPVSNFWVWP	974
960	QY	960	VRLNQTIVINDRQVTFSENLSSCTCHKERLPSHSDFLAELRKAPVNVCSITAVCQRTCCDI	1019
975	DB	975	VELNQEAVMWDEVSHQPNPSLURCSSEKIAPPASDFLAHIQKNPVLQDSITAGCLRPEDV	1034
1020	QY	1020	PPFGIQEENFATLKGNLSPDWIKTSHNHLIIYSTAEILFNDSVFTLLPQCGAFVRSQTE	1079
1035	DB	1035	PSFSVQEEELDFTLKGNLSFGWVRQILQKKVSVSVVAEITFDISVYSQLEQEAFFMAQTT	1094
1080	QY	1080	TKVPEPEVDNPLPLIVGSSVGGILLALITAALYKLGPFPRQYKQMWSE	1128
1095	DB	1095	TVLEKYKHNHPFTPLIVGSSIGILLALITAVLYKVGFPRQYKEMWEE	1143

Search completed: June 7, 2004, 17:19:31
Job time : 22.3484 secs